

# Estimating Discrete Games with Many Firms and Many Decisions: An Application to Merger and Product Variety<sup>\*</sup>

Ying Fan<sup>†</sup>

Chenyu Yang<sup>‡</sup>

University of Michigan

University of Maryland

CEPR and NBER

October 17, 2021

## Abstract

We develop a new method to estimate discrete games based on bounds of conditional choice probabilities. The method does not require solving the game and is scalable to models with potentially many firms and many discrete decisions. We apply the method to estimate a model of entry and product variety for the retail craft beer market in California. Using the estimated model, we simulate a counterfactual merger where a large brewery acquires multiple craft breweries. We find that the merger would cause new firms to enter, non-merging incumbents to add products and merging firms to drop products. However, more products are dropped than added, and consumer surplus decreases. Larger markets are likely to see a smaller loss in per-capita consumer surplus, and markets where the merging firms have a greater market power see a larger reduction in the number of products.

JEL: D43, L13, L41, L66

Keywords: discrete games, incomplete models, entry, product choice, merger, beer

---

<sup>\*</sup>We thank Zibin Huang, Sueyoul Kim, and Xinlu Yao for their excellent research assistance and participants at Caltech, FTC Microeconomics Conference, ITAM, John Hopkins, MIT, NYU IO Day, Stanford, UT Berlin, Yale, and Zhejiang University for their insightful comments. We also thank Ryan Lee, Marc Sorini, and Bart Watson for insights into the craft beer industry. Researchers' own analyses calculated (or derived) based in part on (i) retail measurement/consumer data from Nielsen Consumer LLC ("NielsenIQ"); (ii) media data from The Nielsen Company (US), LLC ("Nielsen"); and (iii) marketing databases provided through the respective NielsenIQ and the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ and Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein

<sup>†</sup>Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; ying-fan@umich.edu.

<sup>‡</sup>Department of Economics, University of Maryland, College Park, MD 20740; cyang111@umd.edu.

# 1 Introduction

To understand the effects of mergers, divestitures, or industry policies on market structure, it is often necessary to estimate a discrete game of firm entry or product choice. In this paper, we consider the estimation of such models when there are many firms and/or each firm makes a large set of discrete decisions. An example is that a firm chooses a subset of products from many potential products to sell in a market. In this case, the action of the firm can be described as a long vector of binary decisions regarding each potential product’s entry. The estimation challenge is a “curse of dimensionality”, as the computational burden of solving a game typically increases exponentially with the number of firms and the complexity of firm decisions. We develop a computationally tractable estimation method, and we use the method to study how an acquisition of craft breweries by a large brewery affects firm entry, product variety, pricing, and welfare.

Our method is based on bounds for conditional choice probabilities. Consider a binary action  $a \in \{0, 1\}$ . Under the assumption that any observed action is not dominated, the equilibrium probability of  $a = 1$  is larger than the probability that  $a = 1$  is a dominant strategy and smaller than the probability that  $a = 1$  is not a dominated strategy. These bounds hold when there is no pure-strategy equilibrium, when there are multiple equilibria, under any equilibrium selection rule, and even when the selection rules differ across markets. More importantly, these bounds are easy to compute. Under the common assumption of additive shocks, such bounds are reduced to cumulative distribution functions evaluated at certain cutoffs. In this paper, we lay out the assumption on firm equilibrium behaviors, explain our bounds, provide step-by-step details on estimation and inference, and use Monte Carlo experiments to evaluate the performance of the method.

We apply the method to consider the effect of a merger on product variety and firm entry. In antitrust litigation, merging parties often argue that entry can potentially mitigate the market power created by a merger. One assumption behind this argument is that the incumbent firms do not change their product offerings. We ask the following questions: does a merger cause incumbents to add or drop products? Does entry occur? What is the overall impact of product adjustments and firm entry on welfare? Do the changes of product variety offset the negative price effects on consumer welfare? How do all these effects vary across markets?

The craft beer industry provides an ideal empirical context to study the effects of merger on the entry and product variety of multi-product firms. Craft breweries have recently become popular targets of acquisitions, attracting the concerns of the antitrust regulators (Codog, 2018). Moreover, there are rich demographic variations across geographical markets

that help to identify consumer tastes and incentives of entry and product adjustments. We focus on the state of California, which has the largest number of craft breweries and the highest craft beer production among the US states according to the Brewers Association, a trade group of the craft beer industry.

To address our research questions, we set up a model to describe consumer demand and firm decisions in the retail beer market in California. The demand side is a flexible random coefficient discrete choice model where we allow for both observed and unobserved heterogeneity in consumer tastes. The supply side is a static two-stage game. In the first stage, each firm is endowed with a set of potential products, and chooses the set of products to sell in a market. A firm can choose an empty set, indicating no entry. In the second stage, firms observe shocks to demand and marginal costs and choose prices simultaneously.

We use a newly compiled data set to estimate our model. Our main data sources are Nielsen Retail Scanner Data and Nielsen Consumer Panel from 2010 to 2016. We supplement these data with information on whether a beer is considered craft based on the designation from the Brewers Association. We further augment the data by hand-collecting the owner and brewery identities and the location of the brewery for each beer.

To estimate the demand model, we combine standard macro moments with a new set of micro-moments based on the panel structure of the consumer survey data. For example, to identify the dispersion in the unobservable heterogeneity in consumer tastes for craft products, we exploit the following intuition: if the standard deviation is large, then a household's taste for craft products is highly correlated over time, implying a large expected total purchase of craft beers in a year conditional on a household ever purchasing a craft beer in the year. The demand estimates reveal substantial unobserved heterogeneity in consumer taste, and that there is little substitution between craft and non-craft products. With the estimated demand and an assumption of Nash-Bertrand pricing, we invert the marginal costs of beers based on the first-order conditions.

We apply our estimation strategy based on dominant strategies to estimate the fixed cost of product entry. This method is well suited in our empirical setting, which features many firms and many potential products. The estimated demand and marginal costs allow us to compute the conditions under which selling a product in a market is a dominant or a dominated strategy, which then give us bounds on the probability that a potential product is observed in a market. We use the inference procedure in Chernozhukov, Chetverikov and Kato (2019) to construct confidence sets. We find higher entry costs for products by independent craft breweries, and that both the mean fixed cost of entry and the variance of the fixed cost shock increase in the sizes of the markets.

Using the estimated model, we simulate the effects of a counterfactual merger where the

largest macro brewery acquires three large craft breweries in 2016. We find that the merger causes new firm entry, which increases product variety. At the same time, merging firms tend to drop products, while the non-merging incumbents tend to add products. The net changes of variety (from new entry and incumbents’ product adjustment) and the associated welfare impacts are both negative. The magnitudes are heterogeneous across markets, with the largest markets seeing more entries and smaller losses in per-capital consumer surplus. Overall, across markets, new entry occurs, but its positive welfare effect is small relative to the loss of consumer surplus. When we consider potential merger efficiencies in the form of a reduction in fixed costs of the acquired breweries’ products, we find that the efficiency gains reduce but not reverse the consumer welfare loss.

In this paper, we study a setting where firms decide on whether to distribute an existing product to a market, and we do not consider the (potentially far larger) costs of dynamic innovation or product development. We find that even in our setting that might favor product entry, the product variety effect of a merger is either negative or not large enough to overcome the loss of consumer surplus.

**Contributions and Literature Review** Our contributions are two-fold. First, we develop and evaluate a new method for estimating discrete choice games. The method can accommodate multiple equilibria, any equilibrium selection rule, different equilibrium selection rules in different markets, the possibility of no pure-strategy equilibrium, and flexible information assumptions. Moreover, computing these bounds does not require solving for an equilibrium and only involves evaluating one-dimensional cumulative distribution functions, making it scalable to settings with many firms or many decisions per firm.

Our method is similar to Ciliberto and Tamer (2009) in two dimensions but differ in the construction of bounds. Both approaches estimate “incomplete models” that do not specify equilibrium selection rules, and both estimate the distribution of fixed cost shocks.<sup>1</sup> However, Ciliberto and Tamer (2009) construct bounds for the probability that an outcome is an equilibrium, where the lower bound is the probability that the outcome is a unique equilibrium and the upper bound is the probability that the outcome is an equilibrium (but it may not be unique). To compute these bounds, one has to simulate a set of fixed cost draws, and for each draw, solve for all equilibria to verify an outcome’s uniqueness. This can be computationally costly if not prohibitive for a model with many firms and a long vector of decisions for each firm, which is the case in our empirical setting. In contrast, we construct bounds for the entry probability of a single product and these bounds are one-dimensional

---

<sup>1</sup>In the auction literature, Haile and Tamer, 2003 explore an incomplete model based on simple assumptions to bound the distribution of bidder valuations.

cumulative distribution functions evaluated at certain cutoffs. Computing these cutoffs does not require specifying equilibrium selection rules or solving the full game. Our approach therefore is especially applicable to estimating games with many players and/or when the decision of a player is represented as a (potentially very long) vector of binary decisions.

Another approach in the literature of estimating discrete games exploits moment inequalities derived from a necessary equilibrium condition that no firm has an incentive to unilaterally deviate from the observed equilibrium. Such an approach typically relies on a mean-zero assumption on non-structural errors (Ho, 2009; Pakes, Porter, Ho and Ishii, 2015; Wollmann, 2018) or support restrictions (Eizenberg, 2014), and does not estimate the distribution of the structural errors associated with the discrete actions. Our approach estimates this distribution and uses it in the counterfactual simulation.

Overall, we consider our approach complementary to existing papers on estimating discrete games. Our approach is suitable for a setting where solving for equilibria is costly and considering the shocks (unobserved by the researchers but observed by the firms) to discrete actions are potentially important for addressing the research questions of interest.<sup>2</sup>

Our second contribution is to the literature of merger, entry response, and product variety. First, entry defense has long been recognized in policy guidelines and investigated in both theoretical (for example, Spector, 2003; Anderson, Erkal and Piccinin, 2020; Caradonna, Miller and Sheu, 2020) and simulation or empirical studies (for example, Werden and Froeb, 1998; Cabral, 2003; Gandhi, Froeb, Tschantz and Werden, 2008). Different from these papers, which focus on entry of single-product firms, we consider multi-product firms. Therefore, it is possible in our model for a merger to decrease product variety while inducing entry, because the incumbents can reduce product offerings. Second, the US Horizontal Merger Guidelines

---

<sup>2</sup>There are four other alternative estimation approaches. First, one can obtain unique equilibrium with additional assumptions and estimate the model with maximum likelihood (Reiss and Spiller, 1989; Garrido, 2020) or simulated method of moments (Berry, 1992; Li, Mazur, Park, Roberts, Sweeting and Zhang, 2021). Alternatively, Illanes (2017) estimates a dynamic discrete choice problem using a semi-parametric latent variable integration method (Schennach, 2014) to deal with selection in unobservables. This approach also avoids solving a game or an optimization problem, but depends on the availability of certain instruments, can result in relatively wide (and sometimes unbounded) confidence sets of parameters. Third, Fan and Yang (2020) directly make assumptions about the distribution of the unobserved fixed cost shock conditional on the observed equilibrium for their merger simulations. In comparison, the approach in this paper estimates this distribution. Finally, in a recent paper, Wang (2020) proposes a hybrid approach that combines the Ciliberto and Tamer (2009) bounds with probability bounds based on the concept of dominant strategies. The computational burden of such an approach is between our methods and Ciliberto and Tamer (2009).

have started to recognize the roles of product variety in a merger.<sup>3</sup> Recent academic work has sought to empirically quantify how a merger affects product variety and the associated welfare impact (Fan, 2013; Wollmann, 2018; Li, Mazur, Park, Roberts, Sweeting and Zhang, 2021; Fan and Yang, 2020; Garrido, 2020) while disallowing firm entry.<sup>4</sup> In this paper, we jointly study entry response and incumbents’ product adjustment due to a merger, and we quantify the net changes of product variety across markets. We also investigate the heterogeneity in these responses. We show that the market power of the merging firms drives the changes of product variety.

There are two modeling differences between our paper and two recent papers on product repositioning (Li, Mazur, Park, Roberts, Sweeting and Zhang, 2021) and entry (Ciliberto, Murry and Tamer, 2020) in the airline industry. First, a firm in our model makes a vector of binary decisions regarding each of their potential products, but a single binary decision on entry or whether to provide non-stop service. By modeling product variety decisions together with firm entry decisions, we allow for the coexistence of new entries and reduced product variety after a merger. Second, Li, Mazur, Park, Roberts, Sweeting and Zhang (2021) and Ciliberto, Murry and Tamer (2020) assume that firms observe demand and marginal cost shocks as well as fixed cost shocks when the firms make decisions on entry. In other words, they account for selection on unobserved demand and marginal costs as well as on unobserved fixed cost shocks. We allow for only the latter selection and address the selection on demand and marginal costs by including a large number of fixed effects in our demand and marginal cost functions. The remaining unobservables are month-to-month product/market-level transient shocks, and we find it reasonable to assume that firms do not observe them when making product choices. We also show that these transient shocks are small.

The rest of the paper is organized as follows. Section 2 explains our estimation method and presents Monte Carlo simulation results. Section 3 discusses the craft beer market in California and the data. Section 4 presents the empirical model and Section 5 the estimation. The estimation results are presented in Section 6. Section 7 describes the counterfactual designs and results. Finally, we conclude in Section 8.

---

<sup>3</sup>The 2010 Guidelines state, “(t)he Agencies also consider whether a merger is likely to give the merged firm an incentive to cease offering one of the relevant products sold by the merging parties. Reductions in variety following a merger may or may not be anti-competitive. Mergers can lead to the efficient consolidation of products when variety offers little in value to customers. In other cases, a merger may increase variety by encouraging the merged firm to reposition its products to be more differentiated from one another”. In comparison, the 1997 US Merger Guidelines did not explicitly mention the roles of product variety.

<sup>4</sup>In the radio industry, a number of papers (e.g., Berry and Waldfogel, 2001; Sweeting, 2010; Jeziorski, 2015) have studied merger, entry and variety but do not quantify the impact on consumer welfare because radio stations do not set prices to listeners. Seim (2006) and Draganska, Mazzeo and Seim (2009) also study entry with endogenous product choice but in an incomplete information framework.

## 2 Discrete Games and Our Estimation Strategy

There are well-known challenges in estimating discrete games. First, there might be multiple equilibria, and the Maximum Likelihood approach may not apply without explicit equilibrium selection rules.<sup>5</sup> Second, a selection issue may complicate a moment-inequality approach: the distributions of unobservables conditional on observed actions differ across these actions. We have discussed the existing methods in the literature in the literature review part of the Introduction. In this section, we present our method by starting with a simple model to illustrate our bounds. We then explain our estimation strategy for more general models. We conclude this section with a set of Monte Carlo experiments to evaluate the performance of our method.

### 2.1 An Illustrative Model and Our Bounds

To illustrate our bounds, we start with a  $2 \times 2$  model where there are only two firms and each firm makes a single binary decision. We later extend the model to a setting with more firms where each firm makes a vector of binary decisions. In this bivariate model, firms 1 and 2 decide whether to enter market  $m$  according to

$$\begin{aligned} Y_{1m} &= \mathbb{1} [\pi_{1m}(Y_{2m}) - C_{1m} - \zeta_{1m} \geq 0], \\ Y_{2m} &= \mathbb{1} [\pi_{2m}(Y_{1m}) - C_{2m} - \zeta_{2m} \geq 0], \end{aligned} \tag{1}$$

where  $Y_{nm} = 1$  indicates entry by firm  $n$  in market  $m$ . In the model,  $\pi_{nm}(Y_{-nm})$  is a variable profit function that depends on the rival action  $Y_{-nm}$ ,  $C_{nm}$  is the fixed cost of entry, and  $\zeta_{nm}$  is a fixed cost shock that follows distribution  $F_\zeta$ .

We consider the following behavioral assumption that is weaker than the Nash equilibrium:

**Assumption 1.**  $Y_{nm}$  is not a dominated strategy for  $n = 1$  or  $2$ .

In other words, we assume that any observed  $Y_{nm}$  is not dominated. This level-1 rationality assumption implies the following bounds for  $\Pr(Y_{nm} = 1)$ :

$$\begin{aligned} &\Pr(Y_{nm} = 1 \text{ is a dominant strategy}) \\ &\leq \Pr(Y_{nm} = 1) \\ &\leq \Pr(Y_{nm} = 1 \text{ is not a dominated strategy}). \end{aligned} \tag{2}$$

---

<sup>5</sup>Tamer (2003) considers an ML estimator in the presence of multiple equilibria for bivariate games.

Given that  $Y_{nm} = 1$  is a dominant strategy if and only if  $\zeta_{nm} < \min \{\pi_{1m}(0), \pi_{1m}(1)\} - C_{nm}$ , and that  $Y_{nm} = 1$  is not a dominated strategy if and only if  $\zeta_{nm} < \max \{\pi_{1m}(0), \pi_{1m}(1)\} - C_{nm}$ , it follows from (2) that

$$\begin{aligned} & F_{\zeta}(\min \{\pi_{nm}(0), \pi_{nm}(1)\} - C_{nm}) \\ & \leq \Pr(Y_{nm} = 1) \\ & \leq F_{\zeta}(\max \{\pi_{nm}(0), \pi_{nm}(1)\} - C_{nm}). \end{aligned}$$

Under the assumption that the entry of the rival reduces the own profit, the inequality further reduces to

$$F_{\zeta}(\pi_{nm}(1) - C_{nm}) \leq \Pr(Y_{nm} = 1) \leq F_{\zeta}(\pi_{nm}(0) - C_{nm}).$$

Before we discuss the general model and estimation, we first compare the bounds with those in the literature and highlight the advantages and differences of our bounds.

### Comparison to Bounds in the Literature

**Ciliberto and Tamer (2009)** One could assume that outcomes observed in data are pure-strategy Nash equilibria and construct bounds for the probability of observing an outcome  $(Y_{1m}, Y_{2m})$ , denoted by  $\Pr(Y_{1m}, Y_{2m})$ , as follows:

$$\begin{aligned} & \Pr((Y_{1m}, Y_{2m}) \text{ is a unique pure-strategy Nash equilibrium}) \\ & \leq \Pr(Y_{1m}, Y_{2m}) \\ & \leq \Pr((Y_{1m}, Y_{2m}) \text{ is a pure-strategy Nash equilibrium}). \end{aligned} \tag{3}$$

The left graph of Figure 1 gives all possible pure-strategy Nash equilibria in each region of  $(\zeta_{1m}, \zeta_{2m})$ . We use  $\Pr(R)$  to represent the probability that  $(\zeta_{1m}, \zeta_{2m})$  is in region  $R$ . We then have the following bounds for the outcome of  $Y_{1m} = 1$  and  $Y_{2m} = 0$ :

$$\sum_{\ell=1,2,4} \Pr(R_{\ell}) \leq \Pr(Y_{1m} = 1, Y_{2m} = 0) \leq \sum_{\ell=1,2,4,5} \Pr(R_{\ell}),$$

and

$$\Pr(Y_{1m} = 1, Y_{2m} = 1) = \Pr(R_3).$$



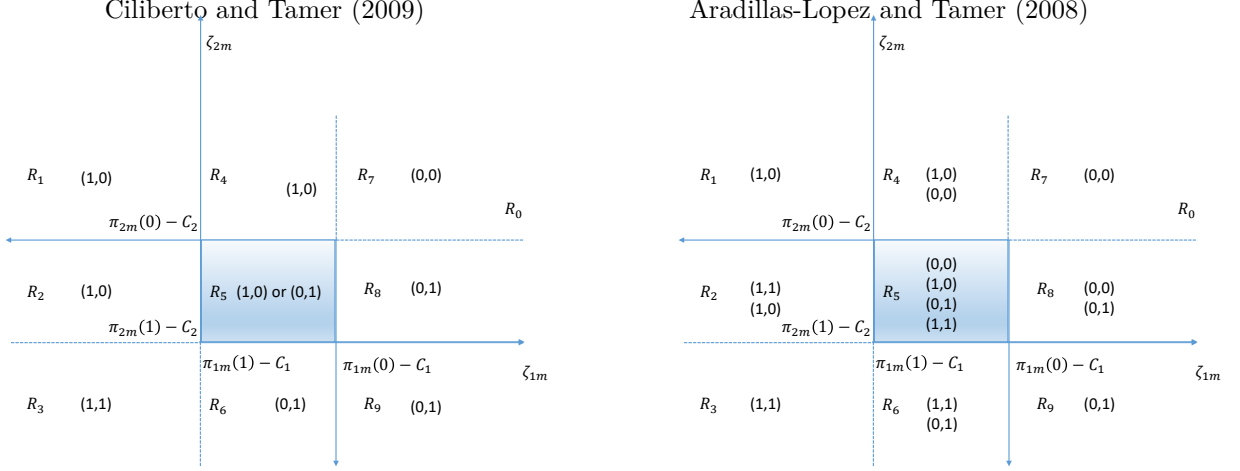


Figure 1: Comparison of Bounds

Therefore, the implied bounds for  $\Pr(Y_{1m} = 1)$  are

$$\sum_{\ell=1}^4 \Pr(R_{\ell}) \leq \Pr(Y_{1m} = 1) \leq \sum_{\ell=1}^5 \Pr(R_{\ell}). \quad (4)$$

In contrast, our bounds are

$$\sum_{\ell=1}^3 \Pr(R_{\ell}) \leq \Pr(Y_{1m} = 1) \leq \sum_{\ell=1}^6 \Pr(R_{\ell}) \quad (5)$$

which are wider than the CT bounds in (4). This is not surprising as the bounds in (4) are derived based on a stronger assumption, i.e., pure-strategy Nash equilibrium, than our level-1 rationality assumption in Assumption 1. However, for a setting with more than two firms and/or each firm makes more than a single binary decision, it quickly becomes intractable to partition the space of shocks and find all equilibria in each region. Instead, one simulates the CT bounds by drawing a large number of the fixed cost shocks  $(\zeta_{1m}, \zeta_{2m})$  and for each of such draws, finding all pure-strategy Nash equilibria by enumerating all possible outcomes and checking whether each one of them is an equilibrium. This procedure could be computationally costly.

**Aradillas-Lopez and Tamer (2008)** Similar to CT, Aradillas-Lopez and Tamer (2008) also construct bounds for the probabilities of equilibrium outcomes but consider assumptions weaker than Nash, such as the level-1 rationality assumption. The right panel of Figure 1 shows all equilibria consistent with the level-1 rationality assumption. Aradillas-Lopez and Tamer (2008) consider the following bounds in the bivariate game:

$$\begin{aligned}
\Pr(R_3) &\leq \Pr(Y_{1m} = 1, Y_{1m} = 1) \leq \sum_{\ell=2,3,5,6} \Pr(R_\ell) \\
\Pr(R_1) &\leq \Pr(Y_{1m} = 1, Y_{1m} = 0) \leq \sum_{\ell=1,2,4,5} \Pr(R_\ell) \\
\Pr(R_9) &\leq \Pr(Y_{1m} = 0, Y_{1m} = 1) \leq \sum_{\ell=5,6,8,9} \Pr(R_\ell) \\
\Pr(R_7) &\leq \Pr(Y_{1m} = 0, Y_{1m} = 0) \leq \sum_{\ell=4,5,7,8} \Pr(R_\ell).
\end{aligned} \tag{6}$$

Despite being derived from the same assumption as our bounds, these inequalities per se do not nest our bounds on a single firm's action: for example, adding up the first two inequalities yields a wider bound on  $\Pr(Y_{1m} = 1)$  than our bounds in (5). The intuition is that the inequalities in (6) focus on the uniqueness of an equilibrium, but do not exploit the uniqueness of actions. For example, although the equilibrium is not unique in  $R_2$ , firm 1 always chooses  $Y_{1m} = 1$  in  $R_2$  because it is the dominant strategy. Furthermore, given that the number of regions to be considered increases exponentially with the number of players,<sup>6</sup> one still has to simulate these bounds.

### Advantages of Our Bounds

There are several advantages of using our bounds for estimation. First, these bounds are one-dimensional CDFs and easy to compute, and computing them does not suffer a dimensionality problem in settings with many firms and/or when each firm makes multiple binary decisions simultaneously (such as product portfolio decisions). Second, the bounds do not rely on equilibrium selection assumptions. Specifically, these bounds hold when there are multiple equilibria, when the equilibrium selection mechanisms differ across markets, or when there is no pure strategy equilibrium for some values of the unobserved fixed cost shocks. Moreover, because our bounds are constructed based on dominant and non-dominated strategies, they are valid under either complete or incomplete information assumptions.<sup>7</sup>

Our bounds are also intuitive. In a single-agent discrete choice model, the inequalities collapse into an equality, and the estimation becomes the GMM estimation of binary choice models (McFadden, 1989). Our approach, therefore, can be considered an extension of the GMM estimation for binary choice models to a game setting.

---

<sup>6</sup>For entry games with  $N$  firms, the number of regions to consider is  $3^N$ .

<sup>7</sup>Grieco (2014) and Magnolfi and Roncoroni (2021) consider estimation of entry games under flexible information assumptions.

## 2.2 General Models and Estimation Using Our Bounds

We consider a market  $m$  with  $N$  firms, where each firm  $n$  makes a vector of binary decisions  $Y_{nm}$  in each market  $m$ . For example, firms decide whether to enter a market and if so, which subset of products from a potential set of products to sell. In this setting,  $Y_{nm} = (Y_{jm}, j \in \mathcal{J}_n)$ , where  $\mathcal{J}_n$  is the set of potential products, and  $Y_{jm} \in \{0, 1\}$  indicates whether product  $j$  is in market  $m$ . Define  $\mathcal{J} = \cup_n \mathcal{J}_n$ . We use  $Y_m = (Y_{jm}, j \in \mathcal{J})$  to denote all firms decisions in the market. We use  $\pi_n(Y_m, X_{nm})$  to denote the variable profit function of firm  $n$ , which depends on its own decisions and its rivals' decisions in the market and a set of firm/market- and/or market-specific observable covariates  $X_{nm}$ . This function is known or has been estimated. For example, it can be the variable profit function at a Nash-Bertrand price equilibrium given the separately estimated demand and marginal costs. We assume that there is a cost associated with choosing  $Y_{jm} = 1$ . This cost is  $c(W_{jm}, \theta) + \zeta_{jm}$ , where  $W_{jm}$  is a set of exogenous covariates and  $\theta$  is a vector of parameters.<sup>8</sup> In the product choice example, the cost represents the fixed cost of entry. Moreover, we assume that the unobserved cost shock  $\zeta_{jm}$  is i.i.d. and follows the distribution  $F_\zeta(\cdot, \sigma_\zeta)$ , where  $\sigma_\zeta$  represents distributional parameters to be estimated.

We define the change in a firm's variable profit when  $Y_{jm}$  is turned from 0 to 1 as

$$\Delta_j(Y_{-jm}, X_{nm}) = \pi_n(Y_{jm} = 1, Y_{-jm}, X_{nm}) - \pi_n(Y_{jm} = 0, Y_{-jm}, X_{nm}), \quad (7)$$

where  $Y_{-jm} = (Y_{j'm}, j' \in \mathcal{J}, j' \neq j)$  represents all but  $j$ 's product entry outcome in market  $m$ . Given the discrete nature of  $Y_{-jm}$ , the following minimum and maximum changes of variable profits exist:  $\underline{\Delta}_j(X_{nm}) = \min_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$  and  $\overline{\Delta}_j(X_{nm}) = \max_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$ .

Following the discussion in the previous section, we can see that the bounds of the conditional probability  $Y_{jm} = 1$  given  $X_{nm}$  and  $W_{jm}$  are

$$\begin{aligned} F_\zeta(\underline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta) &\leq \Pr(Y_{jm} = 1 | X_{nm}, W_{jm}) \\ &\leq F_\zeta(\overline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta). \end{aligned} \quad (8)$$

We define the following moment functions

$$\begin{aligned} L(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) &= F_\zeta(\underline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta) - \mathbb{1}(Y_{jm} = 1), \\ H(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) &= \mathbb{1}(Y_{jm} = 1) - F_\zeta(\overline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta), \end{aligned} \quad (9)$$

---

<sup>8</sup>We implicitly assume that the total cost associated with a vector of binary decisions is the sum of the cost associated with each binary decision. In Appendix D, we consider a model where the total cost is not additively separable across decisions.

The inequalities in (8) imply the following conditional moment inequalities:

$$\begin{aligned} E(L_j(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) | X_{nm}, W_{jm}) &\leq 0, \\ E(H_j(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) | X_{nm}, W_{jm}) &\leq 0. \end{aligned}$$

Following the literature on inference based on conditional moment inequalities, we transform these conditional moment inequalities into unconditional ones as

$$\begin{aligned} E\left(\frac{1}{\#\mathcal{J}} \sum_{j \in \mathcal{J}} L(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) \cdot g^{(k)}(X_{nm}, W_{jm})\right) &\leq 0, \\ E\left(\frac{1}{\#\mathcal{J}} \sum_{j \in \mathcal{J}} H(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) \cdot g^{(k)}(X_{nm}, W_{jm})\right) &\leq 0, \end{aligned} \tag{10}$$

where  $g^{(k)}$ ,  $k = 1, \dots, K$  are non-negative functions of  $(X_{nm}, W_{jm})$  that capture the information in the conditioning variables. In (10), we average over potential products and exploit variations across markets: even under the independence assumption on the fixed cost shock  $\zeta_{jm}$  across  $j$  and  $m$ ,  $Y_{jm}$  across  $j$  within the same market  $m$  may be correlated due to the strategic interdependence among firms, but are independent across markets.

Many papers provide guidance on inferences based on moment inequalities. We give step-by-step details on how we carry out the inference in both Section 2.3 for Monte Carlo experiments and Section 5.2 for our empirical application.

**Identification** The identification of  $(\theta, \sigma_\zeta)$  based on inequalities (8) is similar to the idea of special regressors in entry games (Ciliberto and Tamer, 2009; Lewbel, 2019). We exploit the exogenous variations in  $X_{nm}$  and  $W_{jm}$ . For example, to identify the coefficient in  $\theta$  that corresponds to an indicator variable, we could define a non-negative function  $g^k$  equal to the indicator and compare the entry probability conditional on the indicator being 1 to that conditional on the indicator being 0 (holding other covariates fixed and after appropriate normalization). Similarly, for the coefficient of a continuous variable, we define a set of non-negative indicator functions that indicate whether the value of the continuous variable falls into a certain range. Then, how the entry probability varies across these ranges identifies the corresponding coefficient. The exogenous variations in  $X_{nm}$  and  $W_{jm}$  also allow us to identify the distribution parameters  $\sigma_\zeta$ . For example, if we assume that the distribution of  $\zeta_{jm}$  is fully specified by its variance, then with a large variance, the upper and lower bounds would not co-vary strongly with  $X_{nm}$ : in the special case that  $\zeta_{jm}$  follows a symmetric distribution, both bounds in (7) approach 0.5 as the variance gets arbitrarily large. On the other hand, if the variance is close to 0, both bounds are simultaneously close to 0 if

$\overline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta) < 0$ , or close to 1 if  $\underline{\Delta}_j(X_{nm}) - c(W_{jm}, \theta) > 0$  by the Chebyshev's Inequality. Therefore the model may predict large jumps in entry probabilities even with small changes in the covariates, which can be tested by data.

## 2.3 Monte Carlo Experiments

In this section, we use Monte Carlo experiments to evaluate the performance of our estimation method. We provide step-by-step details about our inference procedure, present the probability that our 95% confidence set covers parameters different from the true parameter values, and compare both the performance and the computational burden of the methods based on our bounds and the CT bounds.

### 2.3.1 Setup

We consider an entry game with  $N$  potential entrants. Each firm  $n$  makes a binary decision  $Y_{nm} \in \{0, 1\}$  where  $Y_{nm} = 1$  represents entering market  $m$ . Firm  $n$ 's post-entry variable profit is

$$\pi_n(Y_{-nm}, X_{nm}) = O_m \cdot \prod_{r \neq n} x_{rnm}^{Y_{rm}},$$

where  $Y_{-nm} = (Y_{rm}, r \neq n)$  represents rival entry decisions,  $O_m > 0$  represents a market-level profit shifter such as the market size, and  $x_{rnm} \in (0, 1)$  represents the competitive impact of firm  $r$ 's entry on firm  $n$ 's profit. We collect covariates in  $X_{nm} = ((x_{rnm}, r \neq n), O_m)$ . As explained in the previous section, the covariates  $X_{nm}$  and the function  $\pi_n(\cdot, \cdot)$  are assumed to be known.

We assume the fixed cost of entry to be  $C + \sigma \zeta_{nm}$ , where the unobservable cost shock  $\zeta_{nm}$  is assumed to a standard normal random variable and i.i.d across firms and markets. The mean fixed cost parameter  $C$  and the standard deviation  $\sigma$  are parameters to be estimated.

In our Monte Carlo experiments, we draw  $O_m$  from a uniform distribution between 0 and 2, and  $x_{rnm}$  from  $a$  to 1. As will be explained later, varying  $a \in (0, 1)$  changes the tightness of our bounds. We set  $C = \sigma = 1$ . For each draw of  $(X_{nm})_{n=1}^N$ , we compute the Nash equilibrium. When there are multiple equilibria, the equilibrium with the highest total profit is selected. For each Monte Carlo experiment, we simulate 500 data sets.

### 2.3.2 Inference

Since the profit function  $\pi_n$  decreases in the entry decision of the rivals, we have

$$\begin{aligned}\min_{Y_{-nm}} \pi_n(Y_{-nm}, X_{nm}) &= \pi_{nm}((1, \dots, 1), X_{nm}) = O_m \cdot \prod_{r \neq n} x_{rnm}, \\ \max_{Y_{-nm}} \pi_n(Y_{-nm}, X_{nm}) &= \pi_{nm}((0, \dots, 0), X_{nm}) = O_m.\end{aligned}$$

We note that being able to compute the extrema of the variable profit functions is important for the computational advantage of our method. These bounds are easy to compute with the profit function in this example or the entry games such as Berry (1992), Seim (2006), Ciliberto and Tamer (2009), Sweeting (2013), and Berry et al. (2016), for more complicated models we provide guides on how to compute these bounds in Section 5.2.

Based on (8), we use the following bounds in estimation:

$$\Phi\left(\left(\left(O_m \cdot \prod_{r \neq n} x_{rnm} - C\right) / \sigma\right)\right) \leq \Pr(Y_{nm} = 1 | X_{nm}) \leq \Phi((O_m - C) / \sigma),$$

where  $\Phi(\cdot)$  is the standard normal distribution function. We note that the gap between the lower and the upper bounds depends on the difference between  $O_m \cdot \prod_{r \neq n} x_{rnm}$  and  $O_m$ . Therefore, as we vary  $a$  and thus the range from which we draw  $x_{rnm}$ , we tighten or widen the bounds.

### Moments

The moment functions are

$$L(Y_{nm}, X_{nm}, C, \sigma) = \Phi\left(\left(\left(O_m \cdot \prod_{r \neq n} x_{rnm} - C\right) / \sigma\right)\right) - \mathbb{1}(Y_{nm} = 1)$$

and

$$H(Y_{nm}, X_{nm}, C, \sigma) = \mathbb{1}(Y_{nm} = 1) - \Phi\left(\left(\left(O_m \cdot \prod_{r \neq n} x_{rnm} - C\right) / \sigma\right)\right).$$

The conditional moment inequalities are

$$E[L(Y_{nm}, X_{nm}, C, \sigma) | X_{nm}] \leq 0, E[H(Y_{nm}, X_{nm}, C, \sigma) | X_{nm}] \leq 0,$$

and we transform them into unconditional moment inequalities:

$$\begin{aligned} E \left[ \frac{1}{N} \sum_n L(Y_{nm}, X_{nm}, C, \sigma) \cdot g^{(k)}(X_{nm}) \right] &\leq 0, \\ E \left[ \frac{1}{N} \sum_n H(Y_{nm}, X_{nm}, C, \sigma) \cdot g^{(k)}(X_{nm}) \right] &\leq 0, \end{aligned}$$

where  $g^{(k)}(X_{nm})$  for  $k = 1, \dots, K$  is a nonnegative function of  $X_{nm}$ , defined as

$$g^{(k)}(X_{nm}) = \mathbb{1} \left( O_m \geq b_1, O_m \cdot \prod_{r \neq n} x_{rnm} \geq b_2 \right),$$

where  $(b_1, b_2) = (0.5, 0.5), (0.5, 1), (1, 0.5)$ , or  $(1, 1)$  for  $k = 1, \dots, 4$ .

### Estimator

We use the inference procedure in Chernozhukov, Chetverikov and Kato (2019) (CCK), where its one-step critical value does not require a tuning parameter. We construct the confidence set by inverting the CCK's test. The test statistic is based on the maximum of  $t$ -type statistics corresponding with each of the moments. Specifically, for each market  $m = 1, \dots, M$ , we define the vector

$$Z_m(C, \sigma) = \left( \sum_n L(Y_{nm}, X_{nm}, C, \sigma) \cdot g^{(k)}(X_{nm}), \sum_n H(Y_{nm}, X_{nm}, C, \sigma) \cdot g^{(k)}(X_{nm}) \right)_{k=1}^K,$$

and we use  $Z_{\tilde{k}m}(C, \sigma)$  to denote a component of the vector, where  $\tilde{k} = 1, \dots, 2K$ . We denote the sample mean and sample standard deviation of the sample moment  $\tilde{k}$  as

$$\hat{\mu}_{\tilde{k}}(\theta, \sigma) = \frac{1}{M} \sum_{m=1}^M Z_{\tilde{k}m}(\theta, \sigma) \text{ and } \hat{\sigma}_{\tilde{k}}(\theta, \sigma) = \sqrt{\frac{1}{M} \sum_{m=1}^M (Z_{\tilde{k}m}(\theta, \sigma) - \hat{\mu}_{\tilde{k}}(\theta, \sigma))^2}.$$

The test statistic is given by

$$\max_{1 \leq \tilde{k} \leq 2K} \frac{\sqrt{M} \hat{\mu}_{\tilde{k}}(\theta, \sigma)}{\hat{\sigma}_{\tilde{k}}(\theta, \sigma)}.$$

The critical value at the significance level of  $\alpha$  is

$$\frac{\Phi^{-1}(1 - \alpha/2K)}{\sqrt{1 - \Phi^{-1}(1 - \alpha/2K)^2/M}},$$

where  $\Phi$  is the standard normal distribution function. To construct the confidence set, we can start with a dense grid of parameters and evaluate the test static at each grid point. The

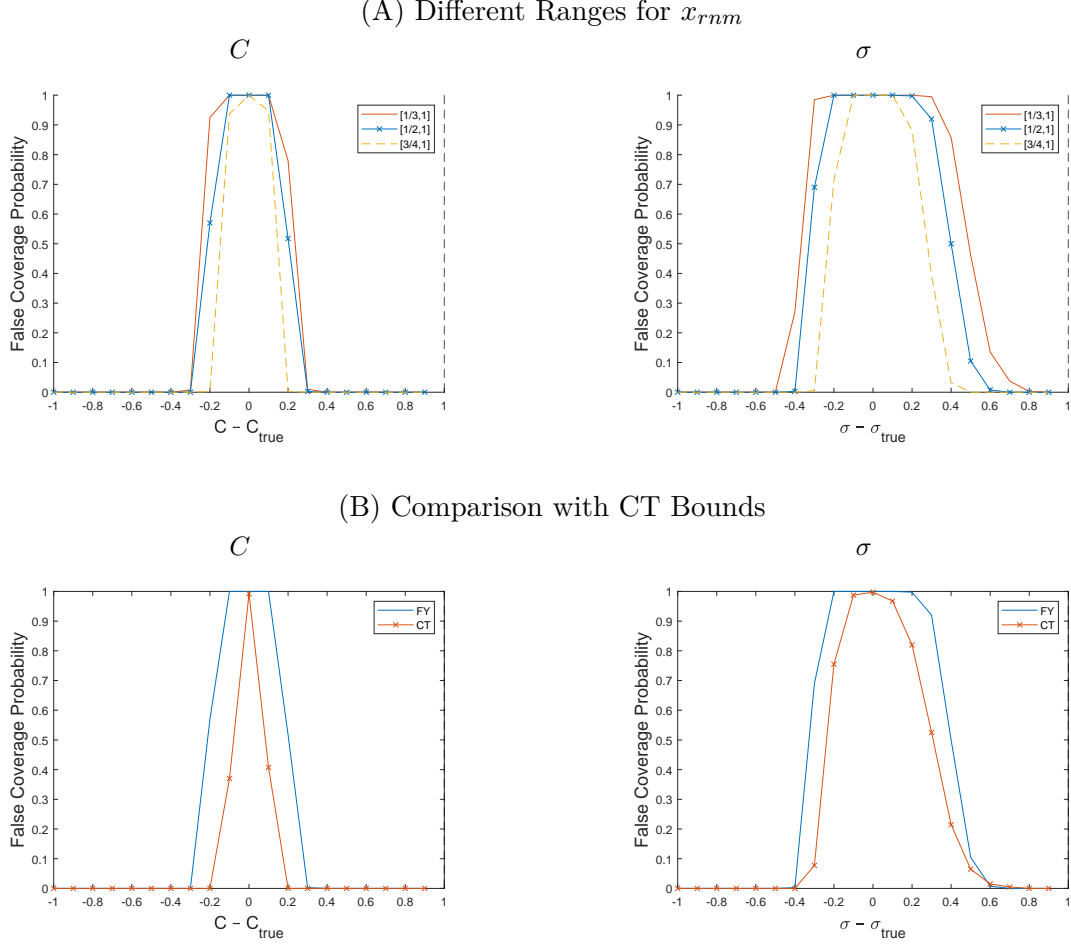


Figure 2: False Coverage Probabilities

confidence set at the 95% confidence level consists of grid points whose test statistic is below the critical value corresponding with  $\alpha = 0.05$ . In practice, we consider 5000 grid points and use each grid point as a starting point to minimize the test statistic. The minimization is stopped when the test statistic falls below the critical value. We collect the resulting set of parameters as the 95% confidence set.

### 2.3.3 Results

We show two sets of results to evaluate the performance of our method. First, we present the coverage probability of our 95% confidence set for the true parameter values (i.e.,  $(C, \sigma) = (1, 1)$ ) and the false coverage probabilities for parameters different from the true parameter values. Second, we compare the statistical power and the computational times of the CCK test statistics based on our bounds and CT bounds.

For the coverage probability of our 95% confidence set for the true parameter values



(i.e.,  $(C, \sigma) = (1, 1)$ ), we simulate 500 data sets with  $N = 2$  and  $M = 2000$ . The coverage probability is the fraction of the 500 data sets where test statistic evaluated at  $(1, 1)$  falls below the 95% confidence level critical value. We repeat this exercise for three different values of  $a$ :  $1/3$ ,  $1/2$  and  $3/4$ , where  $(a, 1)$  is the range for the covariate  $x_{rnm}$ . As explained above, this range determines the gaps between our lower and upper bounds. In each case, we find that the coverage probability is 1. The over-coverage is not surprising because we construct the confidence set for the identified set, and the identified set based on our bounds is likely to be larger than the singleton set of the true parameters.

We also plot the false coverage probabilities of our constructed 95% confidence set for candidate parameter values of  $C$  and  $\sigma$  ranging from 0 to 2. For visibility, we separately compute the false coverage probabilities for  $C$  and  $\sigma$ . For the coverage probability of  $C = 0.8$ , for example, we fix the value of  $C$  at 0.8 and on each of the 500 simulated data set, we minimize the test statistic with respect to  $\sigma$ ; the false coverage probability is computed as the fraction of the data sets where we can find a value of  $\sigma$  so that the test statistic falls below the critical value. We compute these probabilities for candidate parameters of  $C$  from 0 to 2 and analogously for candidate parameters of  $\sigma$  from 0 to 2. The results in the upper two graphs of Figure 2 show that false coverage probabilities are smaller when the covariate  $x_{rnm}$  is drawn from a narrower range, but the probability decreases quickly for parameters further away from the true ones even when  $x_{rnm}$  is drawn from the widest range.

Our second set of results show that our bounds perform slightly worse than CT in terms of power, but comes with a computational advantage. To construct the CCK estimator with the CT bounds, we construct moment inequalities for each possible market outcome  $(0, 0)$ ,  $(1, 0)$ ,  $(0, 1)$ , and  $(1, 1)$  and use the same  $g^{(k)}$  functions given in Section 2.3.2 to make the estimates comparable.<sup>9</sup> The simulated data sets are based on  $a = 1/2$ . We plot the false coverage probability corresponding with each candidate value of  $C$  and  $\sigma$  in the bottom two graphs of Figure 2. The figure shows that the false coverage probabilities using the CT method are moderately smaller than ours (labeled FY).

However, the computational advantage of our bounds grows exponentially with the number of firms in a game. To compare the computational burden of the two methods, we report the time needed to evaluate the test statistic once using our bounds vs. the CT bounds in Table 1.<sup>10</sup> We can see that when  $N = 2$  and  $M = 2000$ , the computational time for evaluating the test statistic once based on the FY bounds is about 1/40th of the time based

<sup>9</sup>To construct the unconditional moments, we interact the moment function corresponding to the lower bound of  $\Pr(Y_{1m}, Y_{2m})$  with  $g^{(k)}(X_{1m})$  as well as  $g^{(k)}(X_{2m})$ . We do the same for the upper bound and for all possible outcomes of  $(Y_{1m}, Y_{2m})$ .

<sup>10</sup>The results are computed on an Intel Core i7-8700K Processor (3.70GHz). We use 100 simulation draws of  $(\zeta_{1m}, \zeta_{2m})$  for simulating the CT bounds.

Table 1: Comparison to CT Bounds: Computational Time

Firms	Market	FY (s)	CT (s)
2	2000	0.0003	0.0139
5	800	0.0003	0.0469
10	400	0.0002	1.3606
15	267	0.0004	43.0739

Notes: the table reports the computational time to evaluate the CCK test statistic corresponding with our bounds (FY) and Ciliberto and Tamer (2009) (CT) bounds.

on the CT bounds, although both methods seem relatively fast when there are only two firms in the model. We note that part of the computation advantage of the FY bounds stems from not having to simulate the bounds for the choice probability. To see how the computational advantage of our bounds becomes more obvious with more firms, we increase the number of firms from 2 to 15. We simultaneously decrease the number of markets to keep the number of firm-market combinations roughly constant to rule out mechanical increases in computational time. From Table 1, we can see that the computational time using the FY bounds is stable as we move from the first row with ( $N = 2, M = 2000$ ) to the last row with ( $N = 15, M = 267$ ). This is not surprising because evaluating our moment functions only involves evaluating a one-dimensional CDF. The computational time using the CT bounds, however, increases from 0.0139 seconds to 43 seconds because the number of possible market outcomes to be checked in order to find all multiple equilibria given one set of draws of  $\xi_{nm}$ ’s increases exponentially with the number of firms.

Overall, the Monte Carlo experiments indicate that our method performs well and is particularly a good alternative to the methods in the literature when the number of players is large.

### 3 Empirical Background and Data

Our empirical analysis focuses on the retail craft beer market in the state of California. According to the 2015 Brewers Association estimates, California accounted for 18% of craft beer volume and 12% of craft breweries in the nation, the highest among all US states. In addition to federal statutes that prohibit “tied-houses”,<sup>11</sup> California passed its own “tied-house” laws and competition laws that further prohibit payments for stocking products (Croxall, 2019). These institutional features motivate our modeling assumption that breweries make the entry and product variety decisions. This simplification also keeps our model tractable.

<sup>11</sup>“Tied-houses” refers to vertical relationships between manufacturers and retailers that exclude small alcoholic beverage makers such as craft breweries from retailers.

Our analysis is based on a newly compiled data set from various sources. First, we obtain data on sales, prices, and product characteristics of beers sold in major retailer chains in the Nielsen data. We use both the aggregate data in the Nielsen Retail Scanner Data and the micro-level panel data in the Nielsen Consumer Panel between 2010 and 2016. We define a product to be a brand (e.g., Samuel Adams Boston Lager) in the Nielsen data. We aggregate the Nielsen data from its original UPC/week level to the product/month level by homogenizing the size of a product (so that a unit is a 12-ounce-12-pack equivalent), adding quantities across weeks within a month, and using the quantity-weighted average price across weeks within a month as the product’s price in that month. Second, we supplement the data with information on whether a beer is considered craft based on the designation by the Brewers Association. Third, we further add hand-collected data on the identities of the corporate owner and the brewery as well as the location of the production facility of each product in our data. For example, Samuel Adams Boston Lager is produced at Samuel Adams Boston Brewery in Boston owned by Boston Beer Company. We define a firm to be a corporate owner (e.g., Boston Beer Company). A firm can own multiple breweries and products. Finally, we merge the data with county demographics from the Census.

We define a market as a retailer-county pair. The Nielsen consumer panel data suggest that cross-retailer shopping appears rare: more than 80% of the households purchased all of their beers from one retailer-county combination in 2016. Similarly, Huang, Ellickson and Lovett (2020) and Illanes and Moshary (2020) find little evidence of retailer competition in the spirit category. In estimation, we define the market size as the average monthly alcohol sales in a market (in the unit of a 12-ounce-12-pack equivalent) times 8, which is the median number of household trips to a retail store in the panel data.<sup>12</sup>

We consider a product was “in” a market in a calendar year if the product’s monthly sales is more than 20 units for more than 6 months in the market in the year. Moreover, for craft products, we keep those by the top 60 craft breweries in the Brewers Association production data sorted by national volume in 2015. We thus focus on breweries established in the 1990s or earlier. In the end, our sample covers 83% of California craft beer quantity in the Nielsen Scanner Data.<sup>13</sup>

We define a firm’s set of potential products in a year as products owned by the firm in any market in the year in our sample. We do not consider the potentially dynamic problem

---

<sup>12</sup>Our results are robust to alternative scaling factors.

<sup>13</sup>Although it is not possible to directly compare the importance of the retail craft beer market with the “on-premise” market (such as taprooms, bars and restaurants) using our data, which cover just the retail segment, the Brewers Association suggests that the retail channel accounts for 65% of the craft volume (Watson, 2016). Probably due to similar data limitations, prior work on the beer industry have also focused on the retail segment (Ashenfelter, Hosken and Weinberg, 2015; Asker, 2016; Miller and Weinberg, 2017; Miller, Sheu and Weinberg, 2019).

Table 2: Annual Total Quantity, Prices and Numbers of Firms and Products

	Total Quantity (12 pk equiv)	Avg. Price (2016 \$)	# Firms Per Year	# Products Per Year
Craft	4,914,209	17	36	135
All	53,465,658	11	54	269

Notes: for each year from 2010 to 2016, we calculate a year’s total and craft quantities, quantity-weighted average prices, number of firms and number of products, and then take the average across years.

Table 3: Shares of Total Quantity and Number of Products by Beer Types

	Ale	Lager	Light
Quantity			
Craft	71.53%	27.13%	0.44%
All	12.50%	46.42%	40.02%
Number of products			
Craft	66.33%	26.87%	0.41%
All	44.19%	39.76%	7.23%

Notes: the shares reflect the proportions of the total quantity of beer from 2010 to 2016, or of the total number of unique products in these years.

of new brewery or brand creation. Instead, we focus on a firm’s decision on whether to sell an existing product to a market, which is arguably far less costly than *de novo* entries.

Table 2 reports summary statistics based on 110 markets present in the data every year from 2010 to 2016. These markets account for 82% of the total quantity from all markets and years. The annual craft beer sales in the sample is, on average, about 5 million units (12-ounce-12-pack), accounting for about 10% of the total beer sales. The average craft beer price is around 17 dollars per unit (in 2016 dollars), which is higher than the average beer price of 11 dollars per unit. Although craft beers account for around 10% of the sales, the number of craft firms and craft products account for over half of the market.

Table 3 provides a breakdown by beer types. Among craft products, ales account for 66% of the product counts and 72% of sales. Lager is another important flavor, with a market share of 27% for craft and around 45% among beers. While light beers account for 46% of all beer sales, their market share within craft products is only 0.44%.

**Fixed Costs** The key primitive in the product variety decisions is the fixed cost of product entry. According to our interviews with industry experts, the main cost of product entry is the cost of the marketing support that a firm needs to provide for a retailer in a local

market.<sup>14</sup> A firm needs to hire sales staff to maintain its relationship with a retailer and consumer awareness of a product. The sales staff’s responsibilities include running local promotions (such as local ads and tasting competitions), and checking with the distributors and the retailers to ensure products are well stocked. Staff salaries and promotional events account for the majority of the expenses. Such costs increase with the number of products and markets.

## 4 Model

### 4.1 Demand

We describe the demand for beer with a random-coefficient discrete-choice model. A product’s characteristics include its flavor types (ale, lager, light, and others),<sup>15</sup> whether it is designated as a craft product, and whether it is imported from outside North America. These types can overlap. For example, Bud Light is a light, non-craft, North American beer, while Samuel Adams Lager is a lager, craft, North American beer. These characteristics of product  $j$  are captured by a vector of indicator variables  $\mathbf{x}_j = (x_j^{\text{ale}}, x_j^{\text{lager}}, x_j^{\text{light}}, x_j^{\text{import}}, x_j^{\text{craft}})$ . We allow both household income and unobserved heterogeneity to affect preferences. We specify the utility function of household  $i$  in market  $m$  from product  $j$  in month  $t$  as

$$\begin{aligned} u_{ijmt} = & (\sigma_0 \nu_i + \kappa_0 y_i) + (\alpha + \kappa_\alpha y_i) p_{jmt} \\ & + \sigma^{\text{ale}} \nu_i^{\text{ale}} x_j^{\text{ale}} + \sigma^{\text{lager}} \nu_i^{\text{lager}} x_j^{\text{lager}} + \sigma^{\text{light}} \nu_i^{\text{light}} x_j^{\text{light}} \\ & + \sigma^{\text{import}} \nu_i^{\text{import}} x_j^{\text{import}} + (\sigma^{\text{craft}} \nu_i^{\text{craft}} + \kappa^{\text{craft}} y_i) x_j^{\text{craft}} \\ & + \beta X_{jm} + F E_j^{\text{demand}} + F E_m^{\text{demand}} + F E_t^{\text{demand}} + \xi_{jmt}, \end{aligned} \quad (11)$$

where  $y_i$  is the logarithm of household  $i$ ’s annual income and  $\nu_i^{(\cdot)}$  is the household-specific unobserved taste shock to each product attribute, which follows a normal distribution and are independent across households. Therefore, the  $\sigma^{(\cdot)}$  parameters capture the dispersion in unobserved household tastes while the  $\kappa^{(\cdot)}$  parameters measure the effect of household income on tastes. The covariates  $X_{jm}$  is a set of indicator functions of whether the distance from the brewery’s nearest production facility to the market falls into a certain distance

---

<sup>14</sup>According to our interviews with the Chief Economist Bart Watson and General Counsel Marc Sorini at the Brewers Association, we think the sunk cost of convincing a retailer to carry the brewery’s products or contracting with a distributor is negligible compared to this flow fixed cost of marketing support. For craft products studied in this paper, it is extremely rare (and highly illegal) for grocers and distributors to charge slotting fees due to the federal and state level alcohol regulations.

<sup>15</sup>Some examples in the category of others include stout, porter and near beers, which collectively account for 0.9% of the craft quantities.

range. Distance potentially plays an important role in the demand of a craft beer because a popular local beer may struggle to gain traction in markets further away due to the lack of name recognition (see, for example, Tamayo, 2009). We also include product fixed effects  $FE_j^{\text{demand}}$ , market fixed effects  $FE_m^{\text{demand}}$ , and month fixed effects  $FE_t^{\text{demand}}$  to capture unobserved factors that vary across at these levels. The error term  $\xi_{jmt}$ , therefore, captures the transient, month-to-month variations of demand shocks specific to a product, market and month combination. We do not include mean coefficients for  $\mathbf{x}_j$  because they are absorbed in product fixed effects. Finally, the last term in (11),  $\varepsilon_{ijmt}$ , is the household idiosyncratic taste, which is assumed to be i.i.d. and follows type-1 extreme value distribution.

This specification gives us the market share  $s_{jmt}(p_{jmt}, p_{-jmt})$  of product  $j$  in month  $t$  and market  $m$ , where  $p_{-jmt}$  is a vector of the prices of all other products in market  $m$  and month  $t$ . Other determinants of demand (product characteristics, fixed effects and demand shocks of all products in the market) are absorbed by the subscript  $jmt$  of the function  $s_{jmt}(\cdot, \cdot)$ . Multiplying the market share by the corresponding market size gives us the demand for product  $j$ ,  $D_{jmt}(p_{jmt}, p_{-jmt})$ .

## 4.2 Supply

The supply side is a two-stage static model. In each market, firms simultaneously choose which beers, if any, to sell. This product choice is made at the beginning of each year  $\tau$  and is fixed throughout the year. We use  $\mathcal{J}_{nm\tau}$  to denote firm  $n$ 's products in market  $m$  in year  $\tau$ . Then, in each month  $t$ , after observing that month's demand and marginal cost shocks, firms simultaneously choose retail prices.

**Stage 2. Pricing** In month  $t$ , firm  $n$  chooses prices  $p_{jmt}$  for all  $j \in \mathcal{J}_{nm\tau}$  to maximize its total variable profit:

$$\max_{p_{jmt}, j \in \mathcal{J}_{nm\tau}} \sum_{j \in \mathcal{J}_{nm\tau}} (p_{jmt} - mc_{jmt}) D_{jmt}(p_{jmt}, p_{-jmt}). \quad (12)$$

The marginal cost  $mc_{jmt}$  is decomposed into a product fixed effect  $FE_j^{\text{mc}}$ , a market fixed effect  $FE_m^{\text{mc}}$ , and a month fixed effect  $FE_t^{\text{mc}}$ , the effect of facility-market distance  $\gamma X_{jm}$  to account for the transportation cost, and a product-market-month specific shock  $\omega_{jmt}$ :

$$mc_{jmt} = FE_j^{\text{mc}} + FE_m^{\text{mc}} + FE_t^{\text{mc}} + \gamma X_{jm} + \omega_{jmt}. \quad (13)$$

In (12), we assume that firms directly set the retail prices. On the technical side, this simplification allows us to avoid excessive computational burdens. In reality, breweries pub-

lish a price menu every few months, and distributors purchase the products and sell them to the retailers. We include product fixed effects, market fixed effects, and month fixed effects in the specification of the brewery marginal cost to capture the distributor and retailer markups in a reduced-form way. Therefore, the underlying assumption is that the markups charged by the distributors and retailers can vary at the product, market and month levels but do not change in counterfactual simulations. Miller and Weinberg (2017) show that a double marginalization model where a brewery first sells to retailers does not significantly change their merger simulation's results.

**Stage 1. Entry and Product Decisions** At the beginning of each year  $\tau$ , each firm  $n$  is endowed with a set of potential products  $\mathcal{J}_{n\tau}$  and decides on its set of products  $\mathcal{J}_{nm\tau}$  in market  $m$  to maximize its expected profit in that market, which is the difference between the expected variable profit  $\pi_{nm}$  and the fixed cost  $C_{nm}$ :

$$\max_{\mathcal{J}_{nm\tau} \subseteq \mathcal{J}_{n\tau}} \pi_{nm}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}) - C_{nm}(\mathcal{J}_{nm\tau}). \quad (14)$$

We now derive the expected variable profit and specify the fixed cost. We first make a timing assumption: when making product decisions, firms observe the product characteristics  $(\mathbf{x}_j, X_{jm})$ , product fixed effects  $(FE_j^{\text{demand}}, FE_j^{\text{mc}})$ , market fixed effects  $(FE_m^{\text{demand}}, FE_m^{\text{mc}})$ , and time fixed effects  $(FE_t^{\text{demand}}, FE_t^{\text{mc}})$  as well as the fixed costs for any all potential products. After firms make product decisions, the month-to-month transient demand and marginal cost shocks  $(\xi_{jmt}, \omega_{jmt})$  are realized in the second stage.

Given this timing assumption, we obtain firm  $n$ 's expected variable profit  $\pi_{nm}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$  by plugging the second-stage equilibrium prices into its profit function, taking the expectation over the transitory demand and marginal cost shocks, and summing over all months in a year. Specifically, we use  $\mathcal{J}_{-nm\tau}$  to denote the set of products that firm  $n$ 's competitors sell in market  $m$  and  $p_{jmt}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$  and  $Q_{jmt}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$  to denote the second-stage equilibrium price and quantity for product  $j$  in the set  $\mathcal{J}_{nm\tau}$  and month  $t$  in year  $\tau$ , which depends on  $(\mathbf{x}_j, X_{jm}, FE_j^{\text{demand}}, FE_j^{\text{mc}}, FE_m^{\text{demand}}, FE_m^{\text{mc}}, FE_t^{\text{demand}}, FE_t^{\text{mc}})$  as well as  $(\xi_{jmt}, \omega_{jmt})$  for all products in market  $m$ . Let  $\xi_{mt} = (\xi_{jmt}, j \in \mathcal{J}_{m\tau} = \mathcal{J}_{nm\tau} \cup \mathcal{J}_{-nm\tau})$  be the collection of the demand shocks for all products in market and define  $\omega_{mt}$  for the marginal cost shocks analogously. Firm  $n$ 's expected variable profit,  $\pi_{nm}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$  in (14) is

$$\begin{aligned} & \pi_{nm}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}) \\ &= \sum_{t=1}^{12} E_{\xi_{mt}, \omega_{mt}} \left( \sum_{j \in \mathcal{J}_{nm\tau}} (p_{jmt}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}) - mc_{jmt}) \cdot Q_{jmt}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}) \right) \end{aligned} \quad (15)$$

The fixed cost function in (14) is specified as

$$C_{nm}(\mathcal{J}_{nm\tau}) = \sum_{j \in \mathcal{J}_{nm\tau}} (W_{jm}\theta + \sigma_m \zeta_{jm\tau}), \quad (16)$$

where  $W_{jm}$  is a vector of covariates, such as whether product  $j$  is a craft product. The fixed cost shock  $\zeta_{jm\tau}$  is assumed to be standard normal and independent across  $j$  and  $m$ , but we allow  $\sigma_m$  to depend on market sizes. This baseline specification of the fixed cost function rules out economies or dis-economies of scope. We extend the model to allow for this possibility and explain how we adapt our estimation strategy in Appendix D.

## 5 Estimation

### 5.1 Estimation of Demand Parameters and Marginal Costs

We combine the aggregate product/market/month-level data of prices, product characteristics, and market shares with the individual/month-level panel data of household purchases to estimate demand parameters. Specifically, we rely on the market share data to identify the mean price coefficient ( $\alpha$ ) and the fixed effects ( $FE_j^{\text{demand}}, FE_m^{\text{demand}}, FE_t^{\text{demand}}$ ). We exploit the panel data and the correlations between household income and beer purchases to identify the standard deviations of the unobservable consumer heterogeneity ( $\sigma^{(\cdot)}$  parameters) and the effect of household income on consumer tastes ( $\kappa^{(\cdot)}$  parameters). We estimate these parameters using the Generalized Method of Moments approach where we combine a set of macro moments and two sets of micro moments.

To deal with the price endogeneity, we use the global barley prices interacted with beer types (i.e., whether a beer is an ale, lager, or light product, and whether it is import or craft) as instrumental variables to construct the macro moments. We choose the prices of barley because barley is a common ingredient in almost all beers, and the interactions with beer types capture the heterogeneity in the proportions of the ingredient in different beers. Figure 3 shows the historical prices of barley in dollars per metric ton, and the price series displays fairly large monthly variations.<sup>16</sup>

We construct a new set of micro moments based on a household's persistence in purchasing decisions to identify the standard deviation parameters  $\sigma^{(\cdot)}$ . For example, if the parameter  $\sigma^{\text{craft}}$  is large, a household's preference for craft products should be highly correlated across months. An implication is that conditional on a household ever purchasing a craft product in a year, the household is likely to have purchased many craft products

---

<sup>16</sup>Data source: <https://fred.stlouisfed.org/series/PBARLUSDM>



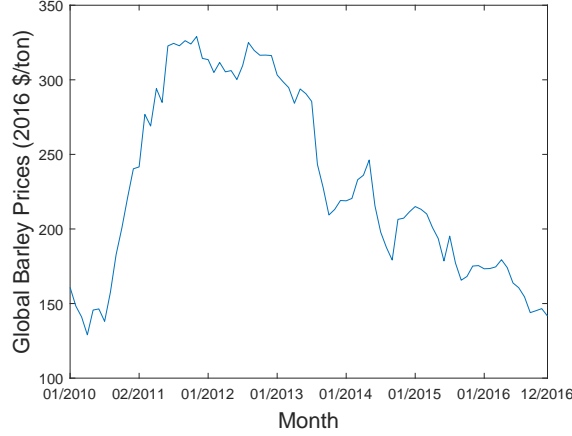


Figure 3: Prices of Barley

throughout the year. More generally, if we use  $q_{it}^f$  to denote a household's purchases of beer with a certain type ( $f \in \{\text{ale, lager, } \dots\}$ ) in month  $t$  in a given year (we suppress the year subscript  $\tau$  here for simplicity), then matching the conditional mean  $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f \geq 1\right)$  helps to identify the parameter  $\sigma^f$ . Similar moments are also useful for identifying the correlation between taste shocks. For example, if households who like type- $f$  products tend to dislike type- $\tilde{f}$  products, then conditional on a household ever purchasing a type- $f$  product, the household should buy few type- $\tilde{f}$  beers throughout the year.

Specifically, we match the model predictions and the empirical counterparts of the following moments (see Appendix A for details on calculation):

- A household's expected annual purchase of a certain type of beer conditional on purchasing at least one unit of this type of beer in the year, i.e.,  $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f \geq 1\right)$ . Matching these moments helps to identify  $\sigma^f$ .
- A household's expected annual purchase of beer conditional on purchasing at least one unit of beer in the year, i.e.,  $E\left(\sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} \geq 1\right)$ , where  $q_{it}$  is a household's total beer purchase in a month. Matching this moment helps to identify  $\sigma_0$ .
- A household's expected annual purchase of a certain type of beer conditional on purchasing at least one unit of craft beer in the year, i.e.,  $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} \geq 1\right)$ . Matching these moments help to identify the taste correlation between craft and type- $f$  beer.

We construct a second set of micro-moments similar to those in Petrin (2002) to identify the income effect on consumer tastes:

- The ratio of average expenditure over average purchase quantity in a year among

households whose income falls into a bin  $\mathcal{I}$ ,<sup>17</sup>

$$E \left( \sum_{t=1}^{12} \text{expenditure}_{it} \mid y_i \in \mathcal{I} \right) / E \left( \sum_{t=1}^{12} q_{it} \mid y_i \in \mathcal{I} \right),$$

where the income bins  $\mathcal{I}$  are  $(0, \$50K]$ ,  $(\$50, \$100K]$  or  $(\$100K, +\infty)$ . Matching these moments helps to identify the income effect on price sensitivity,  $\kappa_1$ .

- $E \left( \sum_{t=1}^{12} q_{it}^{\text{craft}} \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} \geq 1, y_i \in \mathcal{I} \right)$ , which helps to identify  $\kappa^{\text{craft}}$ .
- $E \left( \sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} \geq 1, y_i \in \mathcal{I} \right)$ , which helps to identify  $\kappa_0$ .

The estimation of marginal costs is standard and follows Berry, Levinsohn and Pakes (1995): we back out marginal costs based on the first-order condition of the profit maximization problem in (12).

## 5.2 Estimation of Fixed Cost Parameters

We follow the estimation procedure described in Section 2.2 to estimate the fixed cost parameters, which include the parameters of the mean fixed cost ( $\theta$ ) and the standard deviations of the fixed cost shock ( $\sigma_m$ ). In this section, we explain how we rewrite our empirical model to be consistent with the model described in Section 2.2, provide details on the implementation, and present data patterns that help with the identification of the fixed cost parameters. We can estimate the fixed cost parameters for each year separately and thus suppress the year subscript  $\tau$  in this section for expositional simplicity.

### Reformulate the Model

To be consistent with the model in Section 2.2, we rewrite the profit function in (14)

$$\pi_{nm}(\mathcal{J}_{nm}, \mathcal{J}_{-nm}) - \sum_{j \in \mathcal{J}_{nm}} (W_{jm}\theta + \sigma_m \zeta_{jm}).$$

as

$$\pi_n(Y_{nm}, Y_{-nm}, X_{nm}) - \sum_{j \in \mathcal{J}_n} Y_{jm} (W_{jm}\theta + \sigma_m \zeta_{jm}).$$

---

<sup>17</sup>An alternative moment is the average price  $E \left( \frac{\sum_{t=1}^{12} \text{expenditure}_{it}}{\sum_{t=1}^{12} q_{it}} \mid y_i \in \mathcal{I} \right)$ . This moment is, however, computationally cumbersome because one needs to draw both  $v_i^f$  and  $\varepsilon_{ij}$  to simulate it but only  $v_i^f$  to simulate the moment in the text.

Specifically, we now use a vector of indicators  $Y_{nm} \in \{0, 1\}^{\#\mathcal{J}_n}$  to denote a firm's product portfolio  $\mathcal{J}_{nm} \subset \mathcal{J}_n$ , where  $\mathcal{J}_n$  represents the potential products that firm  $n$  is endowed with. We use  $Y_{jm}$  to denote the element of  $Y_{nm}$  that corresponds to product  $j \in \mathcal{J}_n$ , where  $Y_{jm} = 1$  if  $j \in \mathcal{J}_{nm}$  and 0 otherwise. Therefore, the expected variable profit for given product portfolios of all firms  $\pi_{nm}(\mathcal{J}_{nm}, \mathcal{J}_{-nm})$  can be written as  $\pi_n(Y_{nm}, Y_{-nm}, X_{nm})$ , where the vector  $X_{nm}$  includes all demand and marginal cost covariates (including the fixed effects). Similarly, the total fixed cost  $\sum_{j \in \mathcal{J}_{nm}} (\theta W_{jm} + \zeta_{jm})$  can be written as the summation over all products with  $Y_{jm} = 1$ , i.e.,  $\sum_{j \in \mathcal{J}_n} Y_{jm} (\theta W_{jm} + \zeta_{jm})$ .

In this baseline specification, we assume the additive separability of the fixed costs across products, which is a common assumption in the literature of estimating discrete games. In Appendix D, we extend our method to estimate fixed cost functions that allow for economies or dis-economies of scope (where the additive separability does not hold).

## Details on Estimation

To compute the moment functions, we need to compute  $\underline{\Delta}_j(X_{nm}) = \min_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$  and  $\overline{\Delta}_j(X_{nm}) = \max_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$ , where  $\Delta_j(Y_{-jm}, X_{nm})$  is defined in (7) as the change in firm  $n$ 's expected variable profit when product  $j$  joins the market. Directly solving for the minimum and the maximum of the expected profit over all possible values of  $Y_{-jm}$  is computationally costly because there are  $2^{(\text{length of } Y_{-jm})}$  possible values of  $Y_{-jm}$  and computing  $\Delta_j(Y_{-jm}, X_{nm})$  for each  $Y_{-jm}$  involves solving the stage-2 price game for multiple simulated draws of demand and marginal cost shocks. However, economic intuition suggests that because products are substitutes, we can approximate the minimum and the maximum by, respectively,

$$\underline{\Delta}_j(X_{nm}) \approx \Delta_j((1, \dots, 1), X_{nm}) \text{ and } \overline{\Delta}_j(X_{nm}) \approx \Delta_j((0, \dots, 0), X_{nm}).$$

These approximate extrema are exact for the model in Ciliberto and Tamer (2009), where the variable profit function is a linear function of and is decreasing in the entry decisions of other products. For more general demand and pricing models such as ours, we find that the approximate extrema coincide with the true ones in all of our computational experiments.<sup>18</sup>

We construct the 95% confidence set following Chernozhukov, Chetverikov and Kato (2019). The inference procedure is outlined in Section 2. More details, such as our construc-

---

<sup>18</sup>We randomly sample 100 markets and randomly select  $K$  potential products. For each product  $j$  of the  $K$  selected potential product in a selected market, we hold fixed the entry outcomes of the products not selected, and enumerate all possible outcomes for the other  $K - 1$  products (i.e., all possible  $Y_{-jm}$ ) to find the actual extrema. In all sampled markets for all randomly picked  $K$  products for  $K = 6, 8, 10$ , we find that the approximations coincide with the true ones.

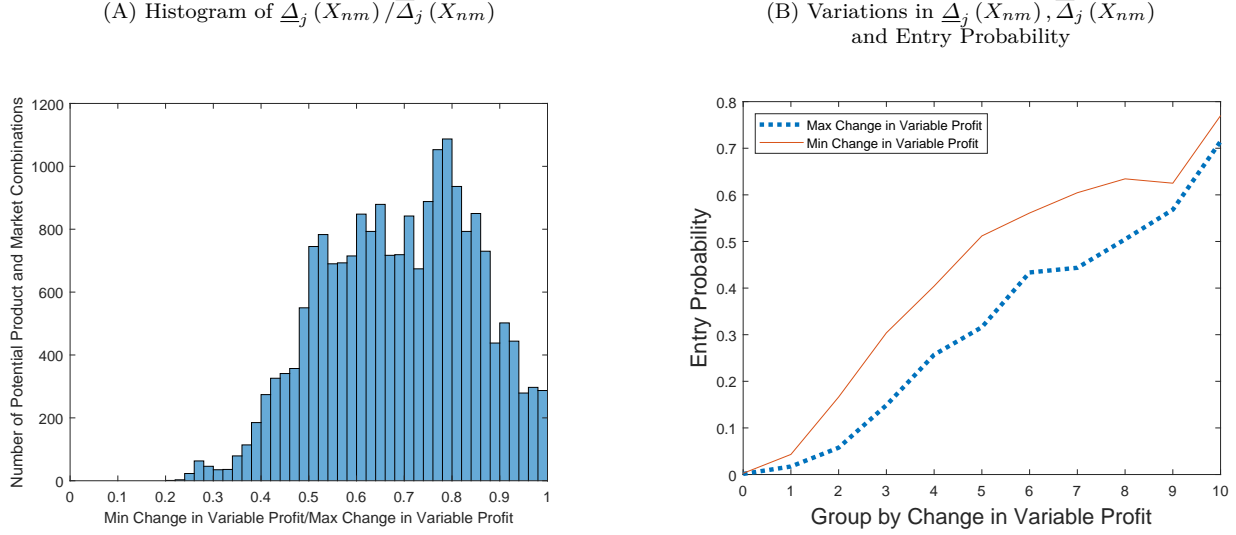


Figure 4: Data Patterns Aiding Identification

tion of the non-negative functions  $g^{(k)}(X_{nm}, W_{jm})$ , are given in Appendix B.

### Data Patterns To Help Identification

We have discussed how our bounds can be informative about the true parameters in Section 2. We now present data patterns in our setting that help with identification. As mentioned, we estimate the fixed cost parameters year by year. We present the results based on the 2016 data.

First, for a large proportion of the observations, the minimum and maximum changes in variable profit, i.e.,  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$ , are relatively close, and as a result, our bounds for the conditional choice probability of these products are fairly tight. We plot the histogram of the ratio  $\underline{\Delta}_j(X_{nm})/\overline{\Delta}_j(X_{nm})$  across all potential product and market combinations in Panel (A) of Figure 4. A larger ratio reflects a smaller difference between the minimum and maximum. The median of the ratio is around 0.7.

Second, there are rich variations in  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$ , and moreover, these variations are informative about variations in entry probabilities. The standard deviations of  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$  are, respectively, \$3950 and \$5527, more than three times of their respective means (\$1178 and \$1714). To see how such variations are associated with variations in the entry probabilities, we discretize  $\underline{\Delta}_j(X_{nm})$  into 10 groups  $[0,240]$ ,  $[240,480]$ , ...,  $[2160,2400]$  and  $[2400, \infty]$ , and for each group  $g$ , we compute the average entry probability

corresponding with observations such that  $\underline{\Delta}_j(X_{nm})$  is in group  $g$  as

$$\frac{\sum_{j,m} \mathbb{1}(\underline{\Delta}_j(X_{nm}) \in g) \cdot Y_{jm}}{\sum_{j,m} \mathbb{1}(\underline{\Delta}_j(X_{nm}) \in g)},$$

where  $Y_{jm} \in \{0, 1\}$  is the entry outcome in the data. In Panel (B) of Figure 4, the red solid line represents these entry probabilities associated with  $\underline{\Delta}_j(X_{nm})$ . We can similarly compute the probabilities corresponding with  $\overline{\Delta}_j(X_{nm})$ , which give us the blue dotted line. The figure shows that the average entry probabilities increase in both  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$ .

What exogenous variations in  $X_{nm}$  generates the variations in  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$  and thus the entry probabilities? In addition to product characteristics, product-, market-, and time-specific fixed effects in the demand and marginal cost functions, variations in market sizes also play an important role. Everything else equal, the returns to entry increase in the size of a market. In Panel (A) of Figure 5, we plot the log of market sizes against the number of craft products, and the figure shows a strong positive correlation between the two.

Another noticeable exogenous variation source is the distance between a production facility and a market. In Panel (B) of Figure 5, we plot the unconditional distribution of distances for all in-state craft potential product/market combinations and the conditional distribution for these combinations where the product enters the market in the data.<sup>19</sup> The conditional distribution has more probability mass at shorter distances, suggesting a negative correlation between distance and entry. We account for these variations in variable profits by including controls for distances in both the demand and marginal cost functions.

## 6 Estimation Results

### 6.1 Demand and Markup

Table 4 reports the demand estimation results. The estimated  $\sigma^{(\cdot)}$  parameters indicate significant heterogeneity in preferences for craft products, imported products, and flavor types. For example, the standard deviation of the unobservable heterogeneity in consumer taste for craft products is 2.44. To translate the estimate into dollar terms, we compare it to the price coefficient of a household with an income of \$50,000, which is  $-2.26 + 0.15 \cdot \ln(\$50,000) = -0.64$ . Therefore, the estimated standard deviation of consumer heterogeneity in taste for craft products is equivalent to a price discount of  $2.44/0.64 = 3.81$  dollars.

The dispersion parameters  $\sigma^{(\cdot)}$  are identified by matching the micro moments capturing

---

<sup>19</sup>Out-of-state craft products tend to be widely distributed and is less affected by distances. Some of the most popular and widely distributed craft products are brewed on or near the east coast.

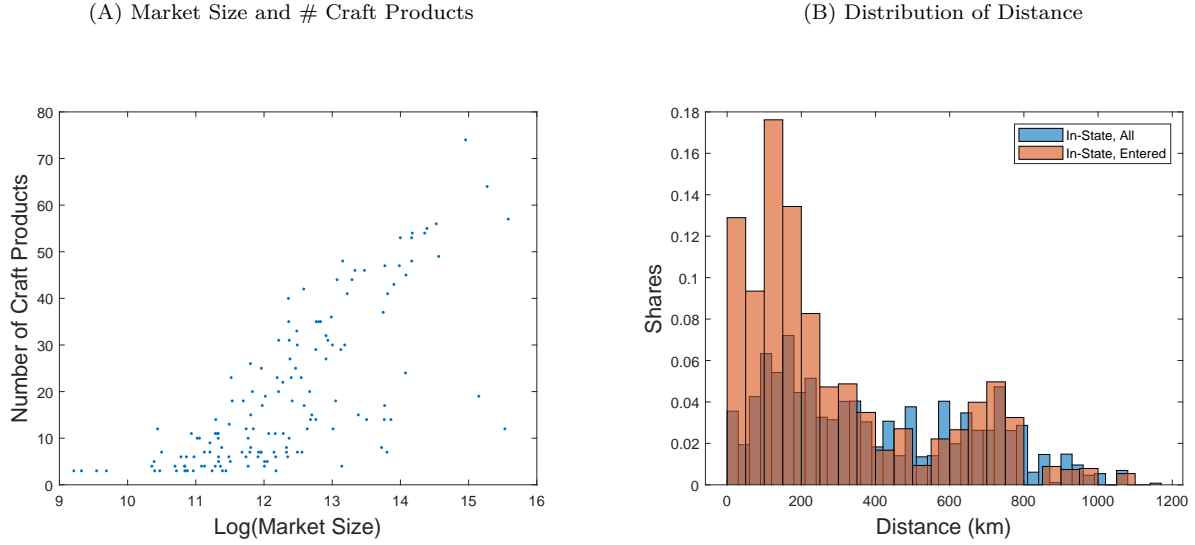


Figure 5: Exogenous Variations Aiding Identification

Table 4: Demand Estimates

Unobserved	$\sigma_0$	0.00	Income Effect	$\kappa_0$	-2.15
Heterogeneity		(0.02)			(0.02)
	$\sigma^{\text{ale}}$	1.98		$\kappa^{\text{craft}}$	1.08
		(<0.01)			(0.02)
	$\sigma^{\text{lager}}$	0.89		$\kappa_\alpha$	0.15
		(<0.01)			(<0.01)
	$\sigma^{\text{light}}$	2.67			
		(<0.01)			
	$\sigma^{\text{import}}$	2.14	Price Coefficient	$\alpha$	-2.26
		(<0.01)			(0.03)
	$\sigma^{\text{craft}}$	2.44			
		(<0.01)			
	$\rho^{\text{craft-light}}$	-0.28			
		(<0.01)			

Note: Standard errors are in parentheses.

Table 5: Micro Moments on Persistence in Purchasing Decisions

		Data	Model
(1)	$E\left(\sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} \geq 1\right)$	7.50	7.80
(2)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{ale}} \mid \sum_{t=1}^{12} q_{it}^{\text{ale}} \geq 1\right)$	3.10	3.91
(3)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{lager}} \mid \sum_{t=1}^{12} q_{it}^{\text{lager}} \geq 1\right)$	5.56	4.17
(4)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{light}} \mid \sum_{t=1}^{12} q_{it}^{\text{light}} \geq 1\right)$	8.03	8.25
(5)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{import}} \mid \sum_{t=1}^{12} q_{it}^{\text{import}} \geq 1\right)$	2.86	2.96
(6)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{craft}} \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} \geq 1\right)$	3.93	4.17

the persistence in a household’s purchasing decisions. Table 5 shows the model fit for these micro moments. For example, the average per-household annual craft purchases among households that purchase at least one unit of craft beers are 3.93. Compared with the unconditional average per-household annual craft purchase of 0.38 units, this micro moment implies that craft beers are purchased by a set of dedicated craft consumers, leading to a significant estimate of  $\sigma^{\text{craft}}$ .

We find that allowing for a correlation between  $\nu_i^{\text{light}}$  and  $\nu_i^{\text{craft}}$  is helpful for matching the conditional purchases of light beers given at least one craft purchase. Summary statistics in Table 3 show that light craft beers count for only 7.23% of the craft beer sales while light beers in general count for 40% of all beer sales. In addition to differences in product characteristics, such a data pattern could suggest that there might be a negative correlation between consumer taste for craft and light beers. Our estimation indeed indicates a negative correlation. The moment  $E\left(\sum_{t=1}^{12} q_{it}^{\text{light}} \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} \geq 1\right)$  is 1.27 in the data and 1.07 according to our estimated model, and it would be 2.14 if such a correlation were not allowed in our model.

As for the effects of income, we find that high-income households are less likely to purchase beer ( $\hat{\kappa}_0 < 0$ ), but they are less price sensitive ( $\hat{\kappa}_\alpha < 0$ ) and have a stronger preference for craft products ( $\hat{\kappa}^{\text{craft}} > 0$ ).

The estimated demand parameters imply that the substitution within craft brands is much larger than the substitution between craft and non-craft products. Table 6 reports the own and cross elasticities among the top-3 non-craft and top-3 craft products in 2016.<sup>20</sup> The elasticities suggest little substitution between the craft products and the non-craft products. This finding is consistent with the histograms of diversion ratios in Figure 6. In this figure, we calculate the diversion ratio for each craft product to other craft products and to non-craft products and plot the histograms. From the figure, we can see that for most craft products,

<sup>20</sup>Per data contract with Nielsen, we refrain from discussing the specific identities of beers or breweries in the data.

Table 6: Elasticities: Top-3 Craft Products and Top-3 Main Products (%)

	Craft			Main		
Craft	-10.09	0.14	0.02	0.01	0.01	0.01
	0.22	-9.52	0.02	0.01	0.01	0.01
	0.04	0.03	-9.16	0.01	0.03	0.01
Main	<0.01	<0.01	<0.01	-5.87	0.04	0.67
	<0.01	<0.01	<0.01	0.08	-6.81	0.08
	<0.01	<0.01	<0.01	0.68	0.04	-5.88

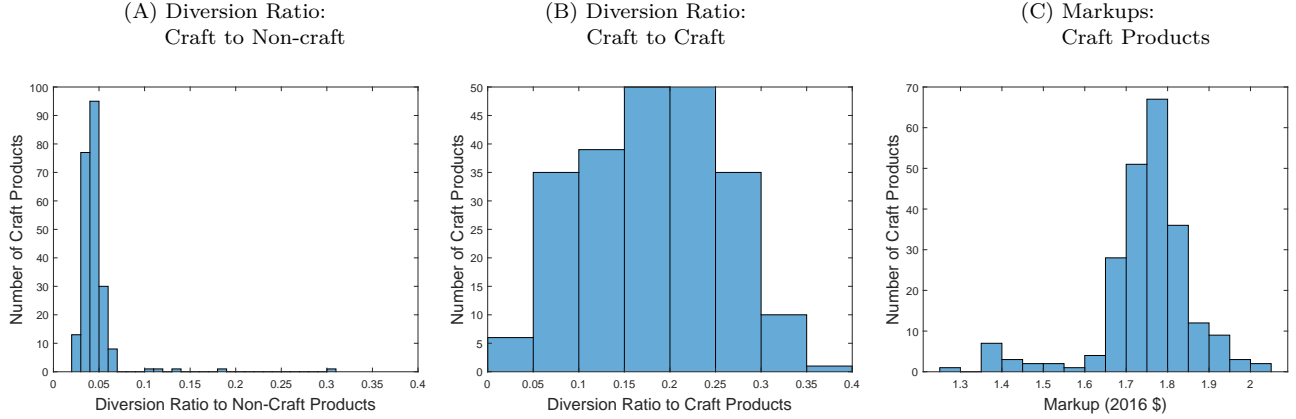


Figure 6: Histograms of Diversion Ratios and Markups

almost no sales would be captured by the non-craft products if the focal craft product's price is increased. In contrast, the distribution of the diversion ratio to other craft products has a mode of around 20%.

We back out marginal costs using the first-order condition in the pricing stage of the game. Panel (C) of Figure 6 shows the distribution of the quantity-weighted markup. The median markup of craft beers is about \$1.7 in 2016 dollars. Some industry sources (e.g., Satran, 2014) put the brewer's margin at 8% of the retail price which, for an average price of \$17, amounts to \$1.4, in line with our estimates.

Finally, we note that observed explanatory variables account for the majority of the variations in the utility levels and marginal costs compared to the month-to-month transient shocks. The  $R^2$ 's from regressing the mean utility and the marginal cost of an observation on observable covariates and fixed effects are both above 0.9, implying a small role for the transient month-to-month shocks. As mentioned, we think it is reasonable, after controlling for product, market and time fixed effects, to assume that firms do not observe the transient shocks when they make entry and product decisions. The finding that these shocks play a small role in explaining demand and marginal cost means that even if the assumption is



violated, the resulting bias is likely to be small.

## 6.2 Fixed Cost

We estimate the fixed cost parameters year by year. The counterfactual simulations in the next section are conducted for the year of 2016. Correspondingly, we report the estimation results using the 2016 data in this section. One unit of observation is a potential product  $j$  and market  $m$  combination. In this part of the estimation, we exclude markets with no craft products. In the end, there are 95 potential products, 149 markets and 14,155 potential product/market combinations in 2016. For each potential product/market combination, we use the demand and marginal cost estimates to compute the moment functions (9). We follow the inference procedures described in Section 2 and the estimation details provided in Section 5, and report the 95% confidence set projected to each parameter in Table 7.

We find a higher fixed cost for independent craft breweries: the 95% confidence set projected to the coefficient of the craft indicator  $\theta_1$  is [\$229, \$1093].<sup>21</sup> We also find that the fixed costs differ by market size. We categorize markets as small, median, and large according to whether the market size is below  $10^5$ , between  $10^5$  and  $5 \times 10^5$ , or above  $5 \times 10^5$  units and allow fixed costs to differ across market-size bins.<sup>22</sup> We find that the fixed costs are higher in larger markets. The standard deviation of the unobservable fixed cost shock also increases in market sizes.

As mentioned, our estimated fixed costs likely reflect the annual expenses on marketing support, and we can compare the estimates with the salary information of the sales personnel from the Brewers Association. A back-of-the-envelope calculation shows that the per-product-retail-county fixed cost, based on a sales representative with a base salary of \$42,000 and responsible for 2 products in 10 retailers in a county, is \$2,100.<sup>23</sup> In comparison, according to our estimation, the 95% confidence interval of the average fixed cost per product (averaged across product/market combinations where the product is sold in the market) is [\$2663, \$3010].

---

<sup>21</sup>We make a distinction between products of independent craft breweries and the products of (former) craft breweries owned by large breweries. We set the craft dummy to 0 for latter products in the fixed cost function, but we still designate these products as craft in demand and marginal cost functions. The underlying assumption is that consumers make purchase decisions based on tastes. Given that the craft beers acquired by large firms are still produced from the same facilities with the same ingredients and procedures, they are likely considered craft by consumers. However, these products may benefit from the distribution and marketing networks of the large firms.

<sup>22</sup>The 25% and 75% quantiles of the market sizes are  $0.84 \times 10^5$  and  $4.5 \times 10^5$  units.

<sup>23</sup>The average annual total wage of a beer sales representative was about \$52,000 in 2018. On average, 20% of the wage comes from commissions. A craft brewery's sales representative handles as few as 10 and as many as more than 50 accounts in a county. An account is typically a retailer in a local region.

Table 7: Estimates of Fixed Costs: Projected 95% Confidence Intervals, 2016

Craft ( $\theta_1$ )	[229.14, 1093.24]
In State $\times$ Craft ( $\theta_2$ )	[-387.82, 208.18]
Market-size specific fixed cost ( $\theta_3$ )	
Small market	[308.95, 938.33]
Medium market	[1027.77, 1468.10]
Large market	[3325.71, 4177.69]
Market-size specific std. dev. ( $\sigma_\zeta$ )	
Small market	[0.00, 522.79]
Medium market	[679.41, 863.25]
Large market	[2511.65, 3424.06]

Note: Estimates in 2016 US dollars.

## 7 Counterfactual Results

### 7.1 Counterfactual Designs

We consider a counterfactual merger where the largest firm in our sample (typically referred to as a “macro brewery”) acquires the three largest craft firms (excluding Boston Beer Company and Sierra Nevada Brewing, which are unlikely merger targets given their sizes) in 2016. There have been mergers during our sample period. Each of these mergers involves a large firm purchasing one small brewery popular in a state. For these mergers, our model predicts small price increases and little to none product variety changes or entry, barring any cost efficiency from these mergers. In contrast, we study a merger of three large craft firms. The merged firm account for 44% of craft sales in 2016. In other words, we are instead concerned with a scenario where this trend of acquisitions continues to a point that the concentration of the craft market approaches the level in the overall beer market.

We allow firms to change their craft products. We hold the non-craft product choices fixed as observed in data to ease the computational burden, but we allow their prices to adjust. The simplification is justified by the estimated small substitution between the craft and non-craft products.

In the simulation, a decision maker is a firm that is observed in any market in our sample. Each firm is endowed with a set of potential products, consisting of the firm’s craft products observed in any market in the 2016 data. In each market, a firm chooses a subset from its potential products. Choosing a non-empty subset implies entry into the market. The merging firms choose products from the union of their potential products. We assume that firms maximize profits in both the product choice and pricing stages. We compute the post-merger product equilibrium using the algorithm in Fan and Yang (2020).

To quantify the effects of the merger and to decompose the overall effects into those due to the price and the product variety adjustments separately, we conduct three counterfactual simulations. Specifically, in the counterfactual simulation described above (CF1), we allow for three adjustment margins—new entry, product adjustment by incumbents, and price adjustment. In CF2, we allow for only the latter two adjustment margins (incumbent product adjustment and price adjustment) by removing products added by new entrants in CF1 and recomputing the pricing equilibrium. In CF3, we allow for only the price effect of the merger by restoring the products in a market to those in the pre-merger market and recomputing the pricing equilibrium. The difference between the outcomes in CF1 and CF3 gives us the overall product variety effect of the merger, which can be further decomposed into the product variety effect due to new entry (CF1 - CF2) and that due to incumbent product adjustment (CF2 - CF3).

For all simulations, we use the point estimates of the demand and marginal cost parameters and sample 30 vectors of fixed cost parameters from their 95% confidence set. We ensure that the sampled parameter vectors include those on the boundary of the confidence set. For each market, each fixed cost parameter vector, and each outcome of interest, we compute the average simulated merger effects across the simulation draws of the demand, marginal cost, and fixed cost shocks. We report the resulting range of this average effect across different fixed cost parameter vectors as the 95% confidence interval of the average effects.

We draw demand and marginal cost shocks directly from their estimated distributions. As for fixed cost shocks, we draw from the estimated distribution while taking into account selection. In other words, our fixed cost draws are consistent with the observed pre-merger equilibrium, which is important to ensure that the pre- and post-merger outcomes are comparable (details on how we draw fixed cost shocks can be found in Appendix C).

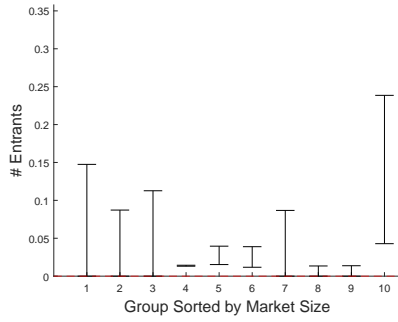
## 7.2 Counterfactual Results

### 7.2.1 Heterogeneous Merger Outcomes Across Markets

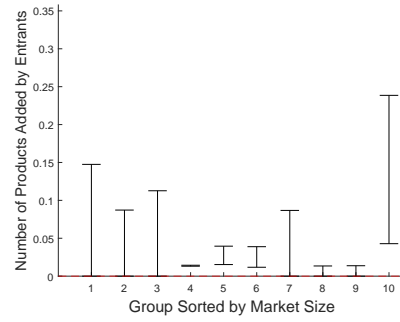
In this section, we show how the effects of the merger vary across markets. For the purpose of presentation, we put markets into 10 groups according to their sizes, where markets in group 1 are the smallest and those in group 10 are the largest. There are 15 markets in groups 1 to 9, and the 14 largest markets are in group 10. Within each group, we average counterfactual outcomes across markets weighted by their sizes.

Figure 7 shows that entry occurs across all market groups, and the largest markets in group 10 see more entries (panel (A)). The number of added products by the new entrants is

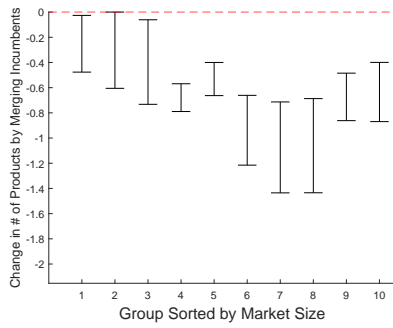
(A) Number of Entrants



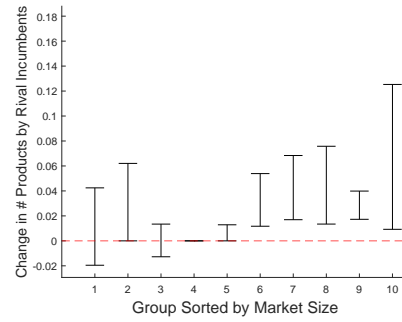
(B) Number of Products Added by Entrants



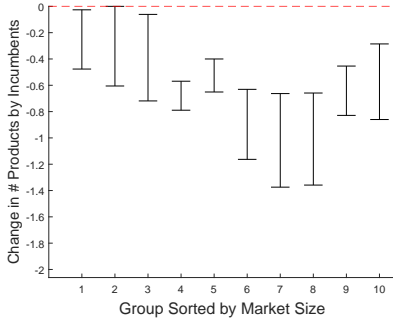
(C) Change in the Number of Products by Merging Firms



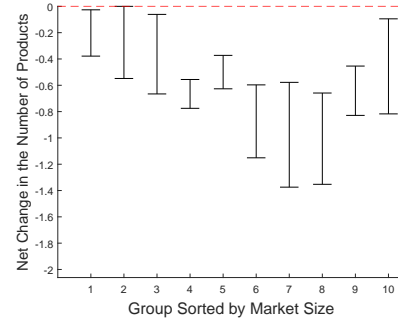
(D) Change in the Number of Products by Non-merging Incumbent Firms



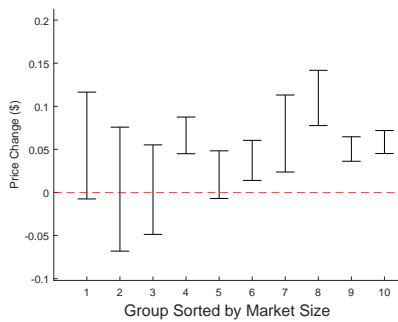
(E) Change in the Number of Products by Incumbents



(F) Change in the Number of Products



(G) Change in Craft Prices



(H) Change in Craft Prices, Variety Fixed

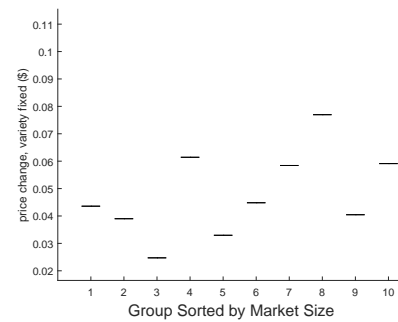


Figure 7: Merger Effects on Entry, Product Variety, and Prices

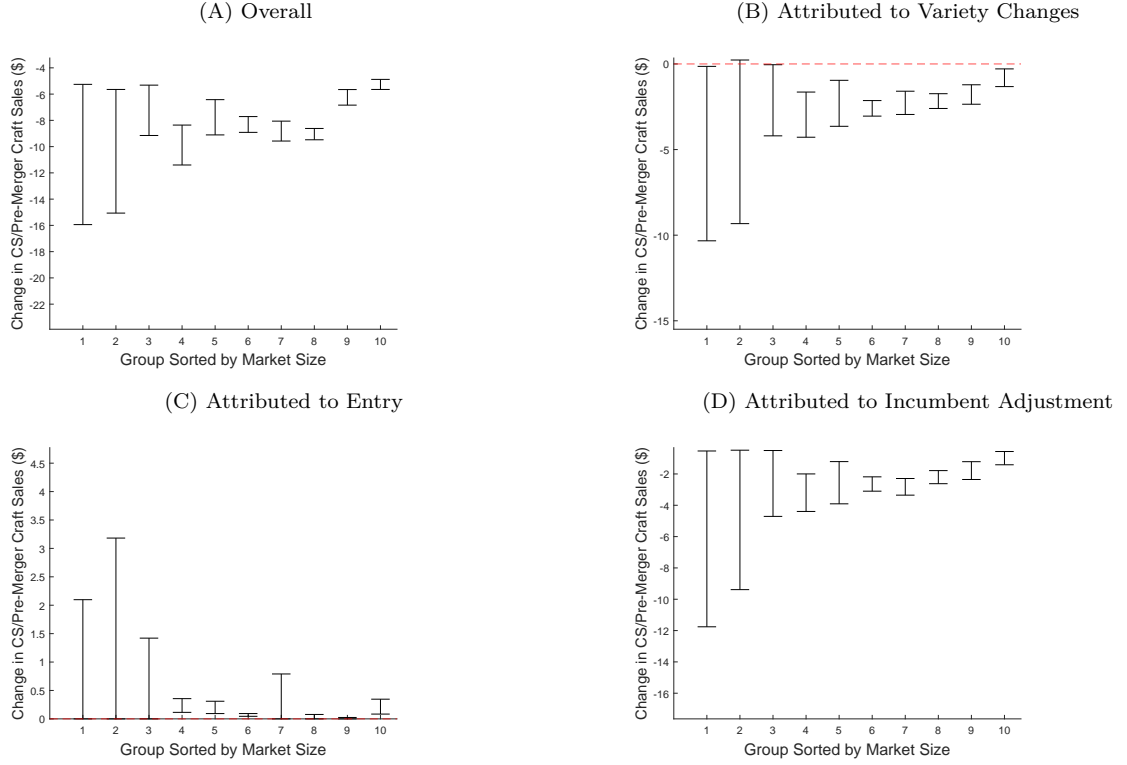


Figure 8: Change in CS/Pre-Merger Craft Sales

almost identical to the number of new entrants (panel (B)), implying that the new entrants, on average, enter with one product. As for the incumbents, the merging firms always drop products (panel (C)) while the non-merging incumbents add products in larger markets (panel (D)). However, the net change in the number of products by the incumbents is always negative (panel (E)). Furthermore, even with the added products by new entrants, the overall number of products decreases (panel (F)). The increase in the quantity-weighted average craft prices is centered around 5 cents, but it could be as large as nearly 15 cents (panel (G)). We also find that the changes of prices when the variety is fixed are similar in magnitudes (panel (H)).

Table 8 summarizes some features of the dropped and added products. The first row suggests that compared to the dropped products, the added products have a bigger pre-merger presence in California. To show this, we first compute the quantile for each potential product in terms of their total sales in the 2016 data among all craft products. Then, for each sampled fixed cost parameter vector and each set of simulation draws of the fixed cost shocks, we find the dropped and added products and take the average of this quantile across products in these two groups separately. Finally, we take the average across the fixed cost simulation draws and report the range across sampled parameter vectors as the confidence

Table 8: Features of Dropped and Added Products

	Dropped	Added
Avg sales quantile	[0.21, 0.30]	[0.43, 0.56]
Avg price quantile	[0.61, 0.68]	[0.38, 0.69]

Notes: the first row reports the average “sales quantiles”, where the sales quantile of a product is defined as its quantile in terms of the total sales in the 2016 data among all craft products, and the average is taken across dropped (first column) or added products (second column) and across simulation draws. The second row reports the average “price quantile”, and the price quantile is defined as a product’s quantile in terms of the quantity-weighted average price (averaged across markets) among craft products. We report the range of these average quantiles across the vectors of fixed cost parameters sampled from their 95% confidence set.

intervals of the average sale quantile for the dropped and added products. The results show that the added products have higher pre-merger sales than the dropped products: the confidence interval of the average pre-merger sales quantile is [43%, 56%] for the added products, compared to [21%, 30%] for the dropped products. We compute similar quantiles for the quantity-weighted average prices to investigate the pre-merger price levels of the added and dropped products. We find that the dropped products are, on average, more expensive than the median: the confidence interval is [0.61, 0.68], above 0.5. The confidence interval for the added products is [0.38, 0.69].

Figure 8 presents the effects of the merger on consumer welfare across markets. We measure the change of the average consumer welfare as the change of the total consumer surplus in a market divided by the market’s pre-merger craft sales. We choose this measure because we find small substitutions between craft and non-craft beers. As a result, the more relevant consumer base for the welfare analysis is the craft consumers. Panel (a) shows that the average loss of craft consumer surplus ranges from 5 dollars to 16 dollars on an annual basis. In panel (B), we break out the average loss attributable to the variety changes. This panel indicates that the decrease in the average craft consumer welfare is smaller in larger markets. A further decomposition of the effects of variety changes indicates that the effect due to new entry is positive (panel (C)) and only partially offsets the negative effect due to the product adjustment by the incumbents (panel (D)), resulting in the net negative change documented in panel (B).

### 7.2.2 Correlation between the Effects of the Merger and Market Characteristics

The heterogeneity of the merger effects on product variety and welfare is likely to be associated with heterogeneity in market characteristics. For example, if the products of the merging firm are close substitutes, the merged firm is more likely to drop products and increase prices significantly. At the same time, pronounced price increases attract new entry and/or product entry by non-merging incumbents. Therefore, a measure of the merging

firms' market power may be positively correlated with both the number of dropped products by the merging firms and the added products by other firms, although the sign is less clear with the net change of the product count. Furthermore, higher fixed costs may be associated with a smaller increase or a larger decline in the number of products. In addition, everything else equal, we expect more entry in larger markets.

We use regressions to document the correlation between the changes of the number of products due to the merger and these market characteristics. In the regressions, one observation is a market. For each market, we measure the merging firms' market power by the increase in the quantity-weighted average price of their products in CF3 (i.e., the variety is fixed). One could consider this measure the realized upward pricing pressure. The average fixed cost of the merging firms in a market is averaged across all of their potential products, where the fixed cost of product  $j$  is given by  $W_{jm}\hat{\theta}$ . In other words, the average fixed cost does not include fixed cost unobservables and thus reflect features of the products and the markets. The average fixed costs of other firms is defined analogously.

Table 9 reports the regression results. We regress the changes in the number of products by merging firms (Column (A)), by other firms (Column (B)), and the net change in a market (Column (C)) on our measure of market power, average fixed costs, and market size. We collect the regression results for each sampled fixed cost parameter vector and report the range of the regression estimates. The estimated coefficients of the variable "variety-fixed price increase" are significant at the 5% level for all sampled parameter values across all three columns. The signs are consistent with the discussion above. Specifically, this variable is negatively correlated with the change in the merging firms' products and positively correlated with that of other firms. The correlation with the net change turns out to be negative. For the other explanatory variables, the estimates are less precise for at least some sampled fixed cost parameter vectors. We also note that the measure of the merging firms' market power accounts for a nontrivial share of the variations in the outcome variables. The range of  $R^2$  values of the three regressions are  $[0.15, 0.58]$ ,  $[0.08, 0.31]$  and  $[0.13, 0.53]$ , respectively. Regressing on just the variety-fixed price increases decreases  $R^2$  only moderately, to  $[0.13, 0.56]$ ,  $[0.03, 0.13]$  and  $[0.08, 0.52]$ .

### 7.2.3 Aggregate Effects

Having established the heterogeneity of the merger effects across markets and documented the correlation between the merger effects and market characteristics, we now turn to the aggregate effects across the simulated 149 markets.

In the left panel for Rows (1)–(9), we report the average changes across markets weighted by the market sizes. In the right panel for Rows (10)–(19), we report the sum across markets.

Table 9: Changes in the Number of Products

	(A) Merging Firms	(B) Other Firms	(C) Market
Avg Price Increase, Merging Firms (\$), Variety Fixed	[-14.64, -4.06]*	[0.33, 0.95]*	[-13.73, -2.88]*
Avg Fixed Cost (\$1000), Merging Firms	[-0.02, 0.00]	[-0.00, 0.00]	[-0.05, 0.00]
Avg Fixed Cost (\$1000), Other Firms	[-0.16, 0.05]	[-0.01, 0.01]	[-0.19, 0.01]
Avg Household Income (\$10,000)	[-0.00, 0.03]	[-0.01, 0.00]	[-0.01, 0.04]
Market Size ( $10^6$ )	[-0.04, 0.08]	[-0.01, 0.04]	[-0.02, 0.21]
$R^2$	[0.15, 0.58]	[0.07, 0.31]	[0.13, 0.53]
$N$	149	149	149

Notes: the dependent variables are the changes in the numbers of products by merging firms (Column (A)) and other firms (Column (B)) and the net change in a market (Column (C)). Each observation is a market. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set. \* indicates significance at the 5% level for all sampled fixed cost parameters.

Table 10: Aggregate Post-Merger Outcomes

Average Change Per Market		Aggregate Change Across Markets	
(1) # of firms	[-2.93, -2.82]	(10) quantity (1000)	[-266.94, -251.57]
(2) # new entrants	[0.02, 0.14]	(11) craft	[-249.76, -230.46]
(3) # of products	[-0.86, -0.33]	(12) craft, merging firms	[-301.64, -283.55]
(4) merging firms	[-0.90, -0.49]	(13) consumer surplus (\$1000)	[-639.00, -602.81]
(5) non-merging incumbents	[0.02, 0.08]	(14) craft beer profits (\$1000)	[97.95, 111.07]
(6) new entrants	[0.02, 0.14]	(15) merging firms	[24.78, 27.68]
(7) average price (\$)	[0.00, 0.00]	(16) total surplus (\$1000)	[-533.03, -504.86]
(8) craft products (\$)	[0.04, 0.07]	$\Delta$ CS decomposition (\$1000)	
(9) craft, merging firms (\$)	[0.13, 0.15]	(17) due to variety change	[-155.43, -106.54]
		(18) due to entry	[6.85, 26.09]
		(19) due to incumbent product adj.	[-164.55, -123.48]

Notes: Rows (1)-(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)-(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.



The 95% confidence interval of the (weighted) average number of new entrants is [0.02, 0.14] (Row (2)). The merged firms drop between 0.5 to 0.9 products (Row (4)), while the rival incumbents add 0.02 to 0.08 products (Row (5)) and new entrants 0.02 to 0.14 products (Row (6)). The average beer price is barely affected. However, the merger increases the average craft prices by about 4 to 7 cents (Row (8)). The average price of the merging firms' craft products increases by about 13 to 15 cents (Row (9)). The drop in the number of products, the increase in prices, and the decrease in quantities (Rows (10)–(12)) lead to a total welfare loss about 0.5 million dollars, aggregated across markets (Row (16)). Allowing for all three adjustment margins, the profit of the merging firms increases by \$24,780 to \$27,680 in 2016 dollars (Row (15)). If product variety is held fixed, the variable profit increase would be \$20,160 (not reported in Table 10), indicating that product adjustments improve the profitability of the merger in equilibrium. On net, incumbent product adjustments exacerbate the loss of consumer surplus by \$106,540 to \$155,430 (Row (19)) while entry recovers \$6,850 to \$26,090 (Row (18)), resulting in a net loss of \$106,540 to \$155,430 dollars of consumer welfare due to product variety changes (Row (17)), which constitutes one quarter to one sixth of the total consumer welfare loss of \$602,810 to \$639,000 dollars (Row (13)).

### 7.3 Merger Efficiency

Mergers involving craft breweries is unlikely to realize efficiency gains in marginal costs. Craft breweries often remain operationally independent after being acquired by a macro brewery, and the beers continue to be brewed at the same facilities. This arrangement stands in contrast to mergers among macro breweries, where the merged firms relocate the production of products and economize on the transportation costs from production facilities to markets (Ashenfelter, Hosken and Weinberg, 2015).

However, craft breweries could benefit from using the marketing networks of the acquirer and reduce the fixed cost of entry. To quantify the equilibrium effects from a fixed cost reduction, we assume that the fixed costs of the acquired craft products decrease by  $\theta_1$ , which is the estimated extra fixed cost faced by independent craft breweries. We report the new aggregate results in Table 11. We find that, on average, the merging firms now add instead of dropping products (Row (4)), but there are fewer new entrants (Row (2)). Overall, the number of products may increase (Row (3)). The merging firms increase the prices by a bigger margin (Row (9)). In the end, the total consumer welfare loss is reduced but not reversed, with a lowered range of \$320,010 to \$512,790 (Row (13)). Unlike the merger efficiency associated with marginal costs, the efficiency associated with fixed costs appears to cause countervailing effects on the number of products: synergies in fixed costs induce

Table 11: Aggregate Post-Merger Outcomes: Merger Efficiencies

Average Change Per Market		Aggregate Change Across Markets	
(1) # of firms	[-2.95, -2.93]	(10) quantity (1000)	[-213.04, -127.44]
(2) # new entrants	[0.00, 0.03]	(11) craft	[-190.51, -99.72]
(3) # of products	[0.00, 1.43]	(12) craft, merging firms	[-229.06, -122.52]
(4) merging firms	[-0.04, 1.42]	(13) consumer surplus (\$1000)	[-512.79, -320.01]
(5) non-merging incumbents	[0.00, 0.01]	(14) craft beer profits (\$1000)	[89.96, 145.94]
(6) new entrants	[0.00, 0.03]	(15) merging firms	[26.77, 103.40]
(7) average price (\$)	[0.00, 0.01]	(16) total surplus (\$1000)	[-422.83, -174.07]
(8) craft products (\$)	[0.07, 0.10]	$\Delta$ CS decomposition (\$1000)	
(9) craft, merging firms (\$)	[0.14, 0.19]	(17) due to variety change	[-16.52, 176.26]
		(18) due to entry	[0.59, 7.25]
		(19) due to incumbent product adj.	[-23.77, 175.67]

Notes: this table reports the results when we take into account potential reductions in fixed costs when a craft brewery is acquired by a macro brewery. Rows (1)-(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)-(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.

new product entry by the merging firms but discourage product entry by new entrants and non-merging incumbents, limiting their overall positive effect on variety and welfare.

## 7.4 (Dis-)economies of Scope

Our current fixed cost specification is additively separable across products and does not allow for economies or dis-economies of scope. We consider the following extension of the fixed cost function:

$$\theta_{0m} \mathbb{1} \left( \sum_{j \in \mathcal{J}_n} Y_{jm} > 0 \right) + \sum_{j \in \mathcal{J}_n} Y_{jm} (W_{jm} \theta + \sigma_m \zeta_{jm}), \quad (17)$$

which is no longer additive in the fixed cost of each product. If  $\theta_{0m} > 0$ , the firm faces a firm-level entry cost in addition to the product-level entry costs, and the fixed cost exhibits economies of scope. Conversely, if  $\theta_{0m} < 0$ , the cost function exhibits dis-economies of scope.

In Appendix D, we derive a new set of inequalities bounding the entry probability of a firm in addition to that of a product. We allow the parameter  $\theta_{0m}$  to be different for the small, medium and large markets, and we find economies of scope in the medium and large markets. Our merger simulation results are, however, robust.

## 8 Conclusion

We develop a new method to estimate discrete games. We apply the method to study the effects of a merger on firm entry, product choices, and prices in the retail craft beer market in California. We make two contributions. On the methodological front, we construct bounds for the probability of a single action instead of a market equilibrium, resulting in an estimator that is easy to compute and scalable to large games (where there are many firms and/or when each firm makes a long vector of binary decisions). On the substantive front, the paper adds to the literature of merger, entry and product variety. For the retail craft beer market in California, we find that entry could happen in most markets after a merger, but its effect is small and does not completely offset the negative effect of the merger. We also find heterogeneity in merger effects across markets. In markets where the merging firms have more market power, there is a larger reduction in product variety. Potential merger efficiencies reduce but are unlikely to reverse the consumer surplus loss.

## References

- Anderson, Simon P, Nisvan Erkal, and Daniel Piccinin (2020), “Aggregative games and oligopoly theory: short-run and long-run analysis.” *The RAND Journal of Economics*, 51, 470–495.
- Aradillas-Lopez, Andres and Elie Tamer (2008), “The identification power of equilibrium in simple games.” *Journal of Business & Economic Statistics*, 26, 261–283.
- Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg (2015), “Efficiencies brewed: pricing and consolidation in the us beer industry.” *The RAND Journal of Economics*, 46, 328–361.
- Asker, John (2016), “Diagnosing foreclosure due to exclusive dealing.” *The Journal of Industrial Economics*, 64, 375–410.
- Berry, Steven, Alon Eizenberg, and Joel Waldfogel (2016), “Optimal product variety in radio markets.” *The RAND Journal of Economics*, 47, 463–497.
- Berry, Steven, James Levinsohn, and Ariel Pakes (1995), “Automobile prices in market equilibrium.” *Econometrica: Journal of the Econometric Society*, 841–890.
- Berry, Steven T (1992), “Estimation of a model of entry in the airline industry.” *Econometrica: Journal of the Econometric Society*, 889–917.

- Berry, Steven T and Joel Waldfogel (2001), “Do mergers increase product variety? evidence from radio broadcasting.” *The Quarterly Journal of Economics*, 116, 1009–1025.
- Cabral, Luis MB (2003), “Horizontal mergers with free-entry: why cost efficiencies may be a weak defense and asset sales a poor remedy.” *International Journal of Industrial Organization*, 21, 607–623.
- Caradonna, Peter, Nathan Miller, and Gloria Sheu (2020), “Mergers, entry, and consumer welfare.” *Available at SSRN 3537135*.
- Chernozhukov, Victor, Denis Chetverikov, and Kengo Kato (2019), “Inference on causal and structural parameters using many moment inequalities.” *The Review of Economic Studies*, 86, 1867–1900.
- Ciliberto, Federico, Charles Murry, and Elie T Tamer (2020), “Market structure and competition in airline markets.” *Available at SSRN 2777820*.
- Ciliberto, Federico and Elie Tamer (2009), “Market structure and multiple equilibria in airline markets.” *Econometrica*, 77, 1791–1828.
- Codog, Aric (2018), “The antitrust roadblock: Preventing consolidation of the craft beer market.” *McGeorge L. Rev.*, 50, 403.
- Croxall, Daniel J (2019), “Helping craft beer maintain and grow market shares with private enforcement of tied-house and false advertising laws.” *Gonz. L. Rev.*, 55, 167.
- Draganska, Michaela, Michael Mazzeo, and Katja Seim (2009), “Beyond plain vanilla: Modeling joint product assortment and pricing decisions.” *QME*, 7, 105–146.
- Eizenberg, Alon (2014), “Upstream innovation and product variety in the us home pc market.” *Review of Economic Studies*, 81, 1003–1045.
- Fan, Ying (2013), “Ownership consolidation and product characteristics: A study of the us daily newspaper market.” *American Economic Review*, 103, 1598–1628.
- Fan, Ying and Chenyu Yang (2020), “Competition, product proliferation, and welfare: A study of the us smartphone market.” *American Economic Journal: Microeconomics*, 12, 99–134.
- Gandhi, Amit, Luke Froeb, Steven Tschantz, and Gregory J Werden (2008), “Post-merger product repositioning.” *The Journal of Industrial Economics*, 56, 49–67.

- Garrido, Francisco Andres (2020), *Mergers between Multi-Product Firms with Endogenous Variety: Theory and an Application to the Ready-To-Eat Cereal Industry*. Ph.D. thesis, Georgetown University.
- Grieco, Paul LE (2014), “Discrete games with flexible information structures: An application to local grocery markets.” *The RAND Journal of Economics*, 45, 303–340.
- Haile, Philip A and Elie Tamer (2003), “Inference with an incomplete model of english auctions.” *Journal of Political Economy*, 111, 1–51.
- Ho, Katherine (2009), “Insurer-provider networks in the medical care market.” *American Economic Review*, 99, 393–430.
- Huang, Yufeng, Paul B Ellickson, and Mitchell J Lovett (2020), “Learning to set prices.”
- Illanes, Gastón (2017), “Switching costs in pension plan choice.” *Unpublished manuscript*.
- Illanes, Gastón and Sarah Moshary (2020), “Market structure and product assortment: Evidence from a natural experiment in liquor licensure.” Technical report, National Bureau of Economic Research.
- Jeziorski, Przemysław (2015), “Empirical model of dynamic merger enforcement—choosing ownership caps in us radio.”
- Lewbel, Arthur (2019), “The identification zoo: Meanings of identification in econometrics.” *Journal of Economic Literature*, 57, 835–903.
- Li, Sophia, Joe Mazur, Yongjoon Park, James Roberts, Andrew Sweeting, and Jun Zhang (2021), “Repositioning and market power after airline mergers.” *RAND Journal of Economics*, *forthcoming*.
- Magnolfi, Lorenzo and Camilla Roncoroni (2021), “Estimation of discrete games with weak assumptions on information.”
- McFadden, Daniel (1989), “A method of simulated moments for estimation of discrete response models without numerical integration.” *Econometrica: Journal of the Econometric Society*, 995–1026.
- Miller, Nathan, Gloria Sheu, and Matthew Weinberg (2019), “Oligopolistic price leadership and mergers: The united states beer industry.” *Available at SSRN 3239248*.
- Miller, Nathan H and Matthew C Weinberg (2017), “Understanding the price effects of the millercoors joint venture.” *Econometrica*, 85, 1763–1791.

- Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii (2015), “Moment inequalities and their application.” *Econometrica*, 83, 315–334.
- Petrin, Amil (2002), “Quantifying the benefits of new products: The case of the minivan.” *Journal of political Economy*, 110, 705–729.
- Reiss, Peter C and Pablo T Spiller (1989), “Competition and entry in small airline markets.” *The Journal of Law and Economics*, 32, S179–S202.
- Satran, Joe (2014), “Here’s how a six-pack of craft beer ends up costing \$12.” URL [https://www.huffpost.com/entry/craft-beer-expensive-cost\\_n\\_5670015](https://www.huffpost.com/entry/craft-beer-expensive-cost_n_5670015).
- Schennach, Susanne M (2014), “Entropic latent variable integration via simulation.” *Econometrica*, 82, 345–385.
- Seim, Katja (2006), “An empirical model of firm entry with endogenous product-type choices.” *The RAND Journal of Economics*, 37, 619–640.
- Spector, David (2003), “Horizontal mergers, entry, and efficiency defences.” *International Journal of Industrial Organization*, 21, 1591–1600.
- Sweeting, Andrew (2010), “The effects of mergers on product positioning: evidence from the music radio industry.” *The RAND Journal of Economics*, 41, 372–397.
- Sweeting, Andrew (2013), “Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry.” *Econometrica*, 81, 1763–1803.
- Tamayo, Andrew (2009), “What’s brewing in the old north state: An analysis of the beer distribution laws regulating north carolina’s craft breweries.” *NCL Rev.*, 88, 2198.
- Tamer, Elie (2003), “Incomplete simultaneous discrete response model with multiple equilibria.” *The Review of Economic Studies*, 70, 147–165.
- Wang, Shuang (2020), “Price competition with endogenous entry: The effects of marriott & starwood merger in texas.”
- Watson, Bart (2016), “On-premise beer data and craft.” <https://www.brewersassociation.org/insights/importance-on-premise-craft-brewers/>.
- Werden, Gregory J and Luke M Froeb (1998), “The entry-inducing effects of horizontal mergers: An exploratory analysis.” *The Journal of Industrial Economics*, 46, 525–543.

Wollmann, Thomas G (2018), “Trucks without bailouts: Equilibrium product characteristics for commercial vehicles.” *American Economic Review*, 108, 1364–1406.

## A Details on Micro-Moments

In this section, we explain how we compute the micro-moment  $E\left(\sum_{t=1}^{12} q_{it}^{\tilde{f}} \mid \sum_{t=1}^{12} q_{it}^f \geq 1\right)$  from data and the model. The calculation for other micro-moments in Section 5.1 is similar. In this moment,  $q_{it}^f$  is household  $i$ ’s quantity of beer with a certain flavor ( $f = \text{ale, lager or light}$ ) or of a certain characteristic ( $f = \text{import or craft}$ ) in month  $t$ , and  $\tilde{f}$  could be the same or a different type.

To compute this moment in the data, we find all household/year combinations where the household purchase at least one unit of type- $f$  beer in the year, i.e.,  $\sum_{t=1}^{12} q_{it}^f \geq 1$ . For each such household, we then compute the annual household purchase of type- $\tilde{f}$  beer, i.e.,  $\sum_{t=1}^{12} q_{it}^{\tilde{f}}$ , which is then averaged across all such household/year combinations.

To compute this moment from the demand model, we let  $s_{jmt}(\boldsymbol{\nu}, y)$  denote the Logit choice probability of product  $j$  when the vector of unobserved tastes and income are  $(\boldsymbol{\nu}, y)$ , and we let  $G_m(\boldsymbol{\nu}, y)$  denote the distribution of  $(\boldsymbol{\nu}, y)$ , which can vary across markets and is thus indexed by  $m$ . We assume that each consumer has 8 opportunities per month to buy beer (where on each trip the consumer buys 1 or 0 products), which is the average per-household number of trips to the stores in the Nielsen Consumer Panel data.<sup>24</sup> Then, the probability that a household buys type- $f$  products in market  $m$  in year  $\tau$  conditional on  $(\boldsymbol{\nu}, y)$  is

$$\rho_{m\tau}^f(\boldsymbol{\nu}, y) = 1 - \prod_{t=1}^{12} \left(1 - \sum_{j \in \mathcal{J}_{m\tau}^f} s_{jmt}(\boldsymbol{\nu}, y)\right)^8,$$

where  $\mathcal{J}_{m\tau}^f$  is the collection of all type- $f$  products in market  $m$  in year  $\tau$ . The conditional expectation of the annual purchase of type- $\tilde{f}$  beers for households in market  $m$  and year  $\tau$  is, therefore,

$$E_{m\tau} \left( \sum_{t=1}^{12} q_{it}^{\tilde{f}} \mid \sum_{t=1}^{12} q_{it}^f \geq 1 \right) = \int_{\boldsymbol{\nu}, y} \frac{\sum_{t=1}^{12} \sum_{j \in \mathcal{J}_{m\tau}^{\tilde{f}}} 8 \cdot s_{jmt}(\boldsymbol{\nu}, y)}{\rho_{m\tau}^f(\boldsymbol{\nu}, y)} dG_m(\boldsymbol{\nu}, y), \quad (18)$$

where  $E_{m\tau}$  means the expectation conditional on a market  $m$  and a year  $\tau$ . To obtain the average across markets, we weigh these conditional means in each market/year combination by the expected number of households who purchase type- $f$  products, which is the product

---

<sup>24</sup>The demand estimates are robust to 5, 6, 7 opportunities of purchase per month.

of the market size and the unconditional probability of purchasing type- $f$  products in a market/year, i.e.,

$$weight_{m\tau} = MktSize_{m\tau} \cdot \int \rho_{m\tau}^f(\nu, y) dG_m(\nu, y).$$

Therefore, the expected purchase of type- $\tilde{f}$  conditional on having at least one purchase of type  $f$  is

$$E\left(\sum_{t=1}^{12} q_{it}^{\tilde{f}} \mid \sum_{t=1}^{12} q_{it}^f \geq 1\right) = \frac{\sum_{m\tau} E_{m\tau}\left(\sum_{t=1}^{12} q_{it}^{\tilde{f}} \mid \sum_{t=1}^{12} q_{it}^f \geq 1\right) \cdot weight_{m\tau}}{\sum_{m\tau} weight_{m\tau}}.$$

## B Details on the Fixed Cost Estimation and Inference

We define our non-negative functions  $g^{(k)}(X_{nm}, W_{jm})$  as functions of  $W_{jm}$  and the extrema of the variable profits  $(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm}))$ , the latter of which are summary statistics of a long vector of covariates capturing market structure and the characteristics of each potential product. Our covariates  $W_{jm}$  are indicator variables and the extrema of the variable profits  $(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm}))$  are continuous variables. To define the functions  $g^{(k)}(X_{nm}, W_{jm})$ , we specify a series of cutoffs for  $\underline{\Delta}_j(X_{nm})$  and  $\overline{\Delta}_j(X_{nm})$ , which are

$$D^\Delta = \{1200, 4800, 8400, 12000\}.$$

We use  $W_{jm\lambda}$  to denote the  $\lambda^{\text{th}}$  component of  $W_{jm}$ ,  $\lambda = 1, \dots, \Lambda$ . The indicator functions are

$$\mathbb{1}\left(\underline{\Delta}_j(X_{nm}) \geq D_{l_1}^\Delta, \overline{\Delta}_j(X_{nm}) \geq D_{l_2}^\Delta\right), \text{ for } l_1 = 1, \dots, 4, l_2 = 1, \dots, 4;$$

$$\mathbb{1}\left(\underline{\Delta}_j(X_{nm}) \geq D_{l_1}^\Delta, W_{jm\lambda}\right), \text{ for } l_1 = 1, \dots, 4, \lambda = 1, \dots, \Lambda$$

$$\mathbb{1}\left(\overline{\Delta}_j(X_{nm}) \geq D_{l_2}^\Delta, W_{jm\lambda}\right), \text{ for } l_2 = 1, \dots, 4, \lambda = 1, \dots, \Lambda$$

Defining the indicator functions in a pairwise fashion helps to avoid having too few non-zero elements in each moment, which could result in noisy estimates of the moment's standard deviation and possibly under or over-coverage of the confidence set.



## C Fixed Cost Simulation Draws Conditional on Observed Equilibrium Outcomes

We draw the fixed costs that are consistent with both the estimated underlying distribution of fixed cost and the pre-merger, observed outcome as a pure-strategy equilibrium. As explained in Section 7, it is important to take into account the latter requirement, which is essentially a selection issue. To obtain one such set of draws in market  $m$ , we proceed with the following steps:

1. For each potential product  $j$  of firm  $n$ , we calculate an upper and a lower bound of the fixed cost shock  $\zeta_{jm}$  as follows. If  $j$  is in the market before the merger, the bound is

$$(-\infty, \Delta_j(X_{nm})),$$

and if product  $j$  is not in the market, the bound is

$$(\Delta_j(X_{nm}), \infty),$$

where we recall that

$$\Delta_j(X_{nm}) = \pi_n(Y_{jm} = 1, Y_{-jm}, X_{nm}) - \pi_n(Y_{jm} = 0, Y_{-jm}, X_{nm}).$$

Here,  $Y_{-jm} = (Y_{j'm}, j' \in \mathcal{J}, j' \neq j)$  represents the observed entry outcomes other than  $j$ 's in market  $m$ .

2. We simulate draws of the fixed cost shocks for firm  $n$  from a truncated normal distribution with the underlying normal distribution parameterized by mean 0 and variance  $\hat{\sigma}_m^2$ , where the variance estimates are dependent on the size of the market  $m$ . The support of the truncated distribution is consistent with the bounds in Step 1. These draws satisfy the necessary conditions for the observed equilibrium.
3. For each draw from Step 2, we check whether firm  $n$ 's best response to  $\mathcal{J}_{-nm}$  is indeed  $\mathcal{J}_{nm}$ . We find each firm's best response by employing the algorithm in Fan and Yang (2020) using the starting points  $\mathcal{J}_{nm}^0 = \emptyset$  and  $\mathcal{J}_{nm}^0 = \mathcal{J}_n$ . If the algorithm converges to  $\mathcal{J}_{nm}$  from both starting points, we keep the set of draws for  $n$ . If at least one of the starting points does not lead to  $\mathcal{J}_{nm}$ , we go back to Step 2 and re-draw the fixed costs.
4. We repeat this process for every firm  $n$ .

## D (Dis-)Economies of Scope

To estimate  $\theta_{0m}$ , we additionally consider bounds for the conditional probability that a firm has at least one product in a market, denoted by  $\Pr(\sum_{j \in \mathcal{J}_n} Y_{jm} > 0 | X_{nm}, W_{nm})$ , where  $W_{nm} = (W_{jm}, j \in \mathcal{J}_n)$  is the collection of the fixed cost covariates for firm  $n$ 's potential products in market  $m$ . Define  $\zeta_{nm} = (\zeta_{jm}, j \in \mathcal{J}_n)$  similarly. Also, let  $Y_{nm} = (Y_{jm}, j \in \mathcal{J}_n)$  be firm  $n$ 's product decision in market  $m$  and  $Y_{-nm}$  be the opponents' decisions. Finally, denote firm  $n$ 's maximum profit from entering the market for given  $(Y_{-nm}, X_{nm}, W_{nm})$  by

$$\Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm}) = \max_{Y_{nm} \text{ s.t. } \sum_{j \in \mathcal{J}_n} Y_{jm} > 0} \pi_n(Y_{nm}, Y_{-nm}, X_{nm}) - \sum_{j \in \mathcal{J}_n} Y_{jm} (W_{jm}\theta + \sigma_m \zeta_{jm}) - \theta_{0m}.$$

Similar to the discussion in Section 2, we define the minimum

$$\underline{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) = \min_{Y_{-nm}} \Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm})$$

and the maximum

$$\bar{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) = \max_{Y_{-nm}} \Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm}).$$

Under the assumption that the observed brewery entry decisions are not dominated, the bounds for  $\Pr(\sum_{j \in \mathcal{J}_n} Y_{jm} > 0 | X_{nm}, W_{nm})$  is

$$\Pr(\underline{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) > 0) \leq \Pr(\sum_{j \in \mathcal{J}_n} Y_{jm} > 0 | X_{nm}, W_{nm}) \leq \Pr(\bar{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) > 0).$$

We note that by the definition of  $\underline{\Delta}_j(X_{nm})$  and  $\bar{\Delta}_j(X_{nm})$ , we have

$$\sum_{j \in \mathcal{J}_n} Y_{jm} \underline{\Delta}_j(X_{nm}) \leq \pi_n(Y_{nm}, Y_{-nm}, X_{nm}) \leq \sum_{j \in \mathcal{J}_n} Y_{jm} \bar{\Delta}_j(X_{nm}).$$

Therefore, we calculate a lower bound of  $\underline{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm})$  by

$$\tilde{\underline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) = \max_{Y_{nm} \text{ s.t. } \sum_{j \in \mathcal{J}_n} Y_{jm} > 0} \left\{ \sum_{j \in \mathcal{J}_n} Y_{jm} (\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}) \right\} - \theta_{0m}.$$

The integer programming problem in the above can be solved quickly given the additive

structure. Specifically, it can be written as

$$\begin{aligned} & \tilde{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) \\ &= -\theta_{0m} + \begin{cases} \sum_{j \in \mathcal{J}_n} \|\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}\|_+ & \text{if } \exists j \text{ s.t. } \underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm} > 0 \\ \max_{j \in \mathcal{J}_n} \{\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}\} & \text{otherwise.} \end{cases} \end{aligned}$$

In other words, we simply need to calculate the values of each  $\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}$ . Then we sum all the positive terms and subtract  $\theta_{0m}$ . If  $\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm} < 0$  for all  $j \in \mathcal{J}_n$ , we calculate the bound as the maximum of  $\underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm} - \theta_{0m}$ . Similarly, the upper bound of  $\bar{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm})$  is given by

$$\begin{aligned} & \bar{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) \\ &= -\theta_{0m} + \begin{cases} \sum_{j \in \mathcal{J}_n} \|\bar{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}\|_+ & \text{if } \exists j \text{ s.t. } \bar{\Delta}_j(X_{nm}) - \theta W_{jm} - \zeta_{jm} > 0 \\ \max_{j \in \mathcal{J}_n} \{\bar{\Delta}_j(X_{nm}) - W_{jm}\theta - \sigma_m \zeta_{jm}\} & \text{otherwise.} \end{cases} \end{aligned}$$

In the end, we use the following as the lower and upper bounds of the entry probability

$$\Pr(\tilde{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) > 0) \leq \Pr(\sum_{j \in \mathcal{J}_n} Y_{jm} > 0 | X_{nm}, W_{nm}) \leq \Pr(\bar{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}) > 0).$$

These bounds do not have analytic expressions and we simulate them by taking 100 draws of  $\zeta_{nm}$ .

We construct unconditional moments similar to those in Section 2. For the non-negative  $g^{(k)}(\cdot)$  functions, we use  $(\underline{\Delta}_j(X_{nm}), \bar{\Delta}_j(X_{nm}))$  of the top three most profitable products by the firm  $n$  and define the indicator functions as in Appendix B. If a firm has only one or two potential products, we set indicator functions corresponding with the unavailable products to be 0.

We also modify the moments associated with an individual product's outcome. The bounds need to take into account  $\theta_{0m}$ :

$$\begin{aligned} & F_\zeta(\zeta_{jm} < \underline{\Delta}_j(X_{nm}) - W_{jm}\theta - \|\theta_{0m}\|_+, \sigma_\zeta) \\ & \leq \Pr(Y_{jm} = 1 | X_{nm}, W_{jm}) \\ & \leq F_\zeta(\zeta_{jm} < \bar{\Delta}_j(X_{nm}) - \theta W_{jm} - \|\theta_{0m}\|_-, \sigma_\zeta). \end{aligned} \tag{19}$$

The same set of  $g$  functions are used to construct the moments associated with individual products.

We combine moments associated with firm and product entry in the estimation. We

Table D.1: Fixed Cost Estimates: Projected 95% Confidence Interval, Allowing for (Dis-)economies of Scope

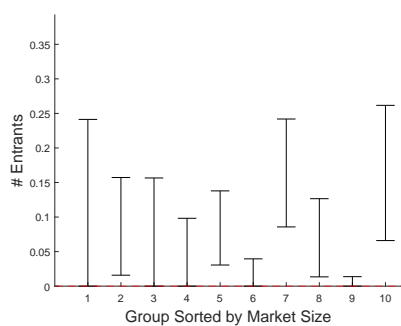
Craft ( $\theta_1$ )	[229.14, 1093.24]	[949.06, 1590.55]
In State $\times$ Craft ( $\theta_2$ )	[-387.82, 208.18]	[-1733.91, -1209.29]
Market-size specific fixed cost ( $\theta_3$ )		
Small market	[308.95, 938.33]	[658.96, 1644.17]
Medium market	[1027.77, 1468.10]	[655.01, 952.30]
Large market	[3325.71, 4177.69]	[1966.51, 3307.71]
Market-size specific std. dev. ( $\sigma_\zeta$ )		
Small market	[0.00, 522.79]	[0.00, 205.29]
Medium market	[679.41, 863.25]	[0.00, 464.35]
Large market	[2511.65, 3424.06]	[1912.77, 2449.15]
Market-size specific firm entry cost (Scope, $\theta_0$ )		
Small market		[-826.28, 277.07]
Medium market		[433.22, 952.24]
Large market		[865.57, 2494.82]

Notes: Estimates in 2016 US dollars.

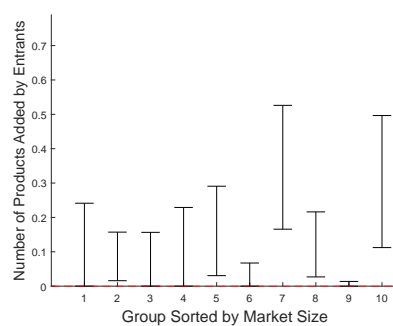
report the 95% projected confidence interval in Table D.1. For comparison, we also copy the results from the baseline specification in the first column. This specification allowing for (dis-)economies of scope estimates a larger cost for craft, but also larger savings for in-state crafts. We find evidence for economies of scope in the medium and large markets.

The counterfactual results on the heterogeneous merger effects in Figures D.1 and D.2 are noisier but still comparable to the baseline results, and we still find that medium-sized markets suffer more per-capita consumer welfare loss than the largest ones. The aggregate results in Table D.2 are also similar to the baseline.

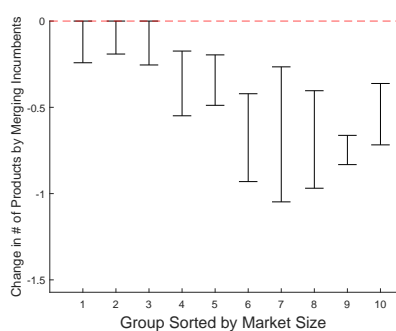
(A) Number of Entrants



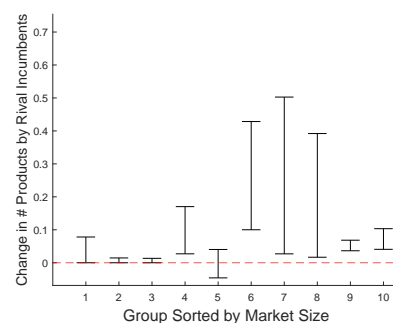
(B) Number of Products Added by Entrants



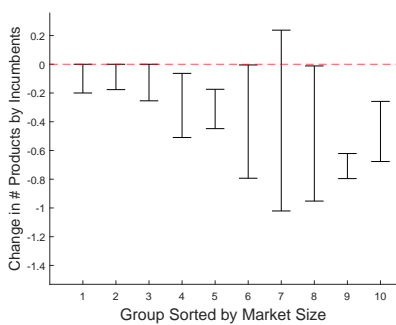
(C) Change in the Number of Products by Merging Firms



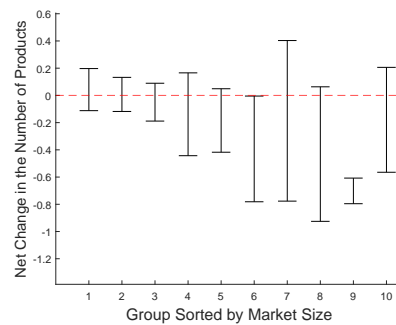
(D) Change in the Number of Products by Non-merging Incumbent Firms



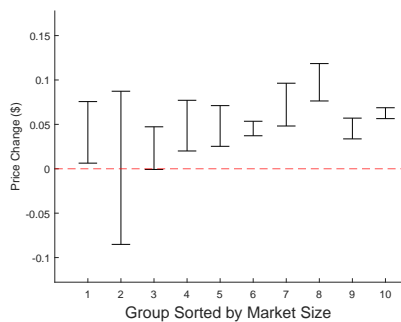
(E) Change in the Number of Products by Incumbent



(F) Change in the Number of Products



(G) Change in Craft Prices



(H) Change in Craft Prices, Variety Fixed

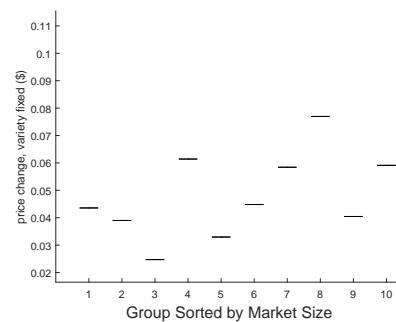


Figure D.1: Product Variety, Entry and Prices

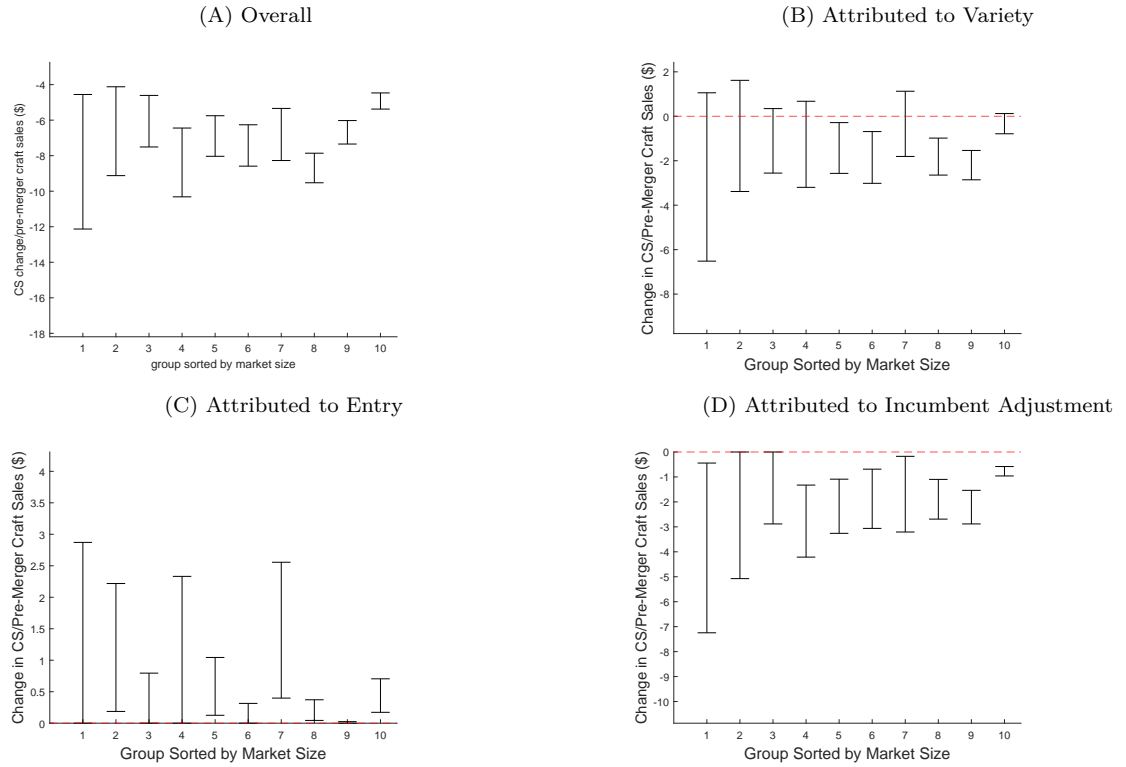


Figure D.2: Change in CS/Pre-Merger Craft Sales

Table D.2: Aggregate Post-Merger Outcomes: Economies of Scope

Average Change Per Market		Aggregate Change Across Markets	
(1) # of firms	[-2.90, -2.79]	(10) quantity (1000)	[-254.19, -225.98]
(2) # new entrants	[0.05, 0.16]	(11) craft	[-232.74, -204.14]
(3) # of products	[-0.54, -0.05]	(12) craft, merging firms	[-292.32, -266.06]
(4) merging firms	[-0.69, -0.48]	(13) consumer surplus (\$1000)	[-606.00, -537.74]
(5) non-merging incumbents	[0.06, 0.15]	(14) craft beer profits (\$1000)	[88.91, 94.92]
(6) new entrants	[0.09, 0.31]	(15) merging firms	[12.41, 18.63]
(7) average price (\$)	[0.00, 0.00]	(16) total surplus (\$1000)	[-511.07, -446.51]
(8) craft products (\$)	[0.05, 0.07]	$\Delta$ CS decomposition (\$1000)	
(9) craft, merging firms (\$)	[0.14, 0.14]	(17) due to variety change	[-109.73, -41.47]
		(18) due to entry	[19.01, 50.59]
		(19) due to incumbent product adj.	[-139.09, -79.06]

Notes: Rows (1)-(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)-(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.